Fisheries and Oceans Canada | Pêches et Océans Canada
Ecosystems and Oceans Science | Sciences des écosystèmes et des océans

Canadian Science Advisory Secretariat (CSAS)

Research Document 2026/006

National Capital Region

Assessment of Port Ocean Prediction System Developed Under Canada's Oceans Protection Plan: Saint John, New Brunswick

Stephanne Taylor[1], Rachel Horwitz[1], Michael Dunphy[2], Maxim Krassovski[2], Hauke Blanken[2], Simon St-Onge Drouin[3], Adam Drozdowski[1]

[1] Fisheries and Oceans Canada
Bedford Institute of Oceanography
1 Challenger Drive
Dartmouth, Nova Scotia B2Y 4A2

[2] Fisheries and Oceans Canada
Institute of Ocean Sciences
9860 West Saanich Road
Sidney, British Columbia V8L 4B2

[3] Fisheries and Oceans Canada
Maurice Lamontagne Institute
850 route de la Mer
Mont-Joli, Quebec G5H 3Z4

February 2026

Canada

Foreword

This series documents the scientific basis for the evaluation of aquatic resources and ecosystems in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/
DFO.CSAS-SCAS.MPO@dfo-mpo.gc.ca

ISSN 1919-5044
ISBN 978-0-660-97522-1 Cat. No. Fs70-5/2026-006E-PDF

Correct citation for this publication:

Taylor, S., Horwitz, R., Dunphy, M., Krassovski, M., Blanken, H., St-Onge Drouin, S., and Drozdowski, A. 2026. Assessment of Port Ocean Prediction System Developed Under Canada's Oceans Protection Plan: Saint John, New Brunswick. DFO Can. Sci. Advis. Sec. Res. Doc. 2026/006. xii + 126 p.

Aussi disponible en français :

Taylor, S., Horwitz, R., Dunphy, M., Krassovski, M., Blanken, H., St-Onge Drouin, S. et Drozdowski, A. 2026. Évaluation des modèles portuaires de prévisions océaniques développés dans le cadre du Plan de Protection des Océans : Saint John, Nouveau-Brunswick. Secr. can. des avis sci. du MPO. Doc. de rech. 2026/006. xiii + 132 p.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Canada's Ocean Protection Plan (OPP) was launched in 2016 to support initiatives aimed at protecting our marine environment from anthropogenic pressures (DFO, 2016). To this end, the oceanography sub-initiative of OPP developed six high-resolution operational port-scale hydrodynamic models, to improve safe navigation, and provide operational emergency response to events such as marine oil spills. The models were downscaled from the Coastal Ice-Ocean Prediction Systems East and West (CIOPS-E, CIOPS-W). Atmospheric forcing was provided by the High-Resolution Deterministic Prediction System (HRDPS).

Model evaluation is an integral part of this development and here the port models are evaluated against available observational data and the parent model using a multiyear hindcast. Evaluations of 48-hour forecasts are performed during a two month period. Additionally, model performance during major storms is examined.

The details of the configuration and evaluation of the port ocean prediction system for Saint John, New Brunswick (POPS-SJ) are presented. The two-level nested system is found to provide improved predictions for water level and sea-surface temperature relative to the existing CIOPS-E system. The hindcast of the POPS-SJ system is also more skillful than CIOPS-E at reproducing observations, including water level, water properties, and especially water velocities.

1 INTRODUCTION

Under the Improving Drift Prediction and Nearshore Modelling (DPNM) sub-initiative of Canada's Oceans Protection Plan (OPP), high-resolution models were developed for six Canadian ports and their approaches to enhance the Government of Canada's ocean modelling capabilities in support of environmental protection and marine safety applications (e.g., drift prediction for oil spills) and safety for navigation via a hydrographic e-navigation application. The six ports (three on the east coast and three on the west coast) were selected as at-risk for environmental incidents owing to their high tanker traffic and complex navigational needs. Models have been developed for the west coast ports of Kitimat, Lower Fraser River and Vancouver Harbour, and for the east coast ports of Saint John, the Strait of Canso, and the St. Lawrence estuary.

The models have been developed with both hindcast and forecast capabilities. For each port, a multi-year hindcast is presented with model validation against observations of water levels, velocities, temperature and salinity, as well as a drift experiment conducted using available drifters. Forecast performance is assessed over a two-month period.

The purpose of this document is to review the performance of the model for the port of Saint John. The design is common to all port models and is presented in Section 2, while Section 3 describes the specifics of the Saint John port model. The evaluation parameters used to evaluate all models are detailed in Section 4, and Sections 5 and 6 respectively analyze the long hindcast evaluation results and the forecasts evaluation results for the port of Saint John. The performance of the Saint John port model is summarized in Section 7, and the main key findings are listed in Section 8.

2 PORT MODEL DESIGN

The port models ("port ocean prediction systems," or POPS) follow a common structure designed to constrain system complexity. We use the same ocean general circulation model and code version, downscaling strategy, preprocessing tools, surface forcing product, and automation suite for all six POPS. The model grids, configuration/tuning, open boundary forcing, and freshwater input sources differ between POPS configurations.

2.1 MODEL SELECTION

The precursor to much of the OPP port modelling effort was conducted using the Finite Volume Community Ocean Model (FVCOM) (Chen, Liu, and Beardsley 2003) under the World Class Tanker Safety System (WCTSS) program, yielding prototype models for ports of Canso, Kitimat, and Vancouver Harbour. Even earlier, there have been modelling efforts based on Backhaus (Backhaus 1983; 1985) yielding models for the St. Lawrence estuary (Saucier and Chassé 2000) and the Gulf of St. Lawrence (Saucier 2003). These were then followed by implementation of an equivalent NEMO model by Environment and Climate Change Canada (ECCC), the Regional Marine Prediction System (RMPS) GSL. The RMPS was operational from 2011 to 2021, though none of these earlier models were considered for use in OPP. The WCTSS FVCOM prototypes used unstructured model meshes to resolve the coastal regions to within tens of metres or less, with resolution gradually decreasing to kilometres to match the parent model grid on the open boundary.

FVCOM has no history of operational deployment in Canada, while NEMO 3.6 (Nucleus of European Modelling of the Ocean) is used operationally by the Canadian Operational Network for Coupled Environmental PredicTion Systems (CONCEPTS). In the early part of OPP, a

comparison exercise between FVCOM 4.1 and NEMO 3.6 (Nudds et al. 2020) informed the decision-making process to select which codebase to use for the POPS models. Both codebases were used to construct models of the Bay of Fundy and the Port of Saint John (Paquin et al. 2020), which was chosen due to the large tides and complex circulation. NEMO 3.6's most significant deficiency in this context is the lack of wetting and drying, so if NEMO 3.6 could provide satisfactory results in the Bay of Fundy, then it would likely be suitable for use in the other regions. The two models were evaluated on how well they matched observations, as well as on computational efficiency, stability, and robustness. Both models were found to be skillful at reproducing observed data: neither model was significantly superior, and the choice of model to use going forward rested more heavily on the other factors.

Ultimately the decision was taken to proceed with NEMO 3.6, despite its structured grid, a somewhat coarser nearshore resolution, and higher demand for computational resources, to facilitate operationalization and align with modelling efforts at ECCC. Additional advantages of NEMO include active development that delivers regular code updates and bug fixes, an international NEMO Consortium group where members steer code development, and a well-established international operational modelling community.

Thus the ocean model used for all port ocean prediction systems is the CONCEPTS code: a fork of NEMO 3.6 (Madec 2016) that has been customized to meet the operational needs of CONCEPTS, for example (Dupont et al. 2015).

2.2 DOWNSCALING STRATEGY

The port models are downscaled solutions driven by larger-scale coastal ocean models currently operational at ECCC: the Coastal Ice-Ocean Prediction Systems East and West (CIOPS-E, CIOPS-W; Paquin et al 2021a, Paquin et al 2021b) have 2-2.5 km resolution, and Salish Sea 500 (~500 m resolution) which is part of the CIOPS-W system. Output from these models forms the boundary conditions for our higher-resolution, smaller-area models. We use two levels of nesting to achieve a resolution fine enough to reach port scale. The nesting is one-way (coarse to fine), so no information is fed back to the larger scale models, allowing the models to run sequentially but otherwise independently of each other. This one-way nesting strategy also enables systematic errors to be corrected at the open boundaries.

We do not employ a dynamic ice model. Instead, we use a NEMO feature called "ice if", which uses input ice fields and the local freezing point to assess where ice cover exists, and in those locations it restores the sea surface temperature to the local freezing point and sets heat fluxes to -4 Wm^{-2} (Madec 2016).

We do not employ data assimilation or spectral nudging; all model runs are free runs.

River discharge data is used where available to supply the most realistic freshwater input to the model, and climatology is used when this is not available. Gauge data is also used in some cases to construct water level boundary conditions.

2.3 SIMULATION SEQUENCING

The port models operate in three configurations: hindcast, pseudo-analysis and forecast. Hindcasts are the most straightforward, using larger-scale model forcing and quality-controlled gauge data to drive the models. The model begins in the past from a cold-start (temperature and salinity interpolated as initial conditions, water at rest) or a hot-start (temperature, salinity, velocity, and sea-surface height interpolated as initial conditions) and reaches a spun-up state after a period of adjustment to the forcing. The model output is considered usable once spun up, and the model can run nearly up to the present in this configuration, provided that forcing data is

available. Pseudo-analysis runs are daily runs that keep the model state caught up to near-real-time, and do not include a direct data-assimilation component. Rather, the state is indirectly driven by data via boundary and surface forcing terms. These runs are used to initialize the first forecast of each day and may use different input than hindcasts depending on what data is available in real-time.

The pseudo-analysis and forecast schedule is chosen to match the schedule of the parent models that we use for forcing. Shortly after 00Z each day, a 24 hour pseudo-analysis simulation runs to catch up the model state to 00Z. This process uses restart files (so no spin up needed) and, where possible, uses gauge data drawn from a near-real-time data feed which receives limited quality control. Following the pseudo-analysis, the POPS generates four forecasts per day, each 48 hours long, which start from 00Z, 06Z, 12Z and 18Z. The daily 00Z forecast starts using the restart file from the daily pseudo-analysis and runs for 48 hours, saving a restart file six hours into the simulation. The 06Z forecast starts from this restart file, also saving a restart file six hours in, and similarly for the 12Z and 18Z forecasts. A schematic of this setup is shown in Figure 1. We focus only on the 00Z forecasts in this evaluation.

While the 00Z forecast simulation is nominally started at 00Z each day, in practice, the initialization of the simulation is delayed as it can not begin until all inputs are available from the larger scale models and the pseudo-analysis completes. At time of writing, this delay is approximately 5 hours, such that upon completion, each forecast simulation has about 43 hours of output that is in the future.

2.4 SURFACE FORCING

Surface forcing is derived from the High-Resolution Deterministic Prediction System (HRDPS) (Milbrandt et al. 2016) that runs operationally at ECCC and provides atmospheric weather forecasts four times per day at 2.5 km resolution. This is the highest-resolution operational atmospheric product available and is chosen to be consistent with the forcing used in CIOPS-E/W. In hindcast and pseudo-analysis mode, we use a time-blended form of the HRDPS forecasts, where hours 06–17 from successive forecasts are combined using weighted averaging to form temporally continuous fields with the same blending schedule as CIOPS-E/W. Additional details are given in the technical documentation for CIOPS-W version 1.5 (Paquin et al. 2022). Time blending is not used for forecasts.

The surface forcing is applied to the NEMO model using the CORE algorithms (Large and Yeager 2004) with modifications by ECCC to (a) read input data from the in-house RPN file format and (b) use the lowest diagnostic level of the atmospheric model rather than the conventional 2 m and 10 m data. Precipitation and sea-level pressure variations are also applied to the surface of the model.

2.5 AUTOMATION SUITE

The hindcast, pseudo-analysis and forecast simulations are all managed using ECCC's Maestro sequencing software. We have constructed a Maestro suite that is based on ECCC's CIOPS-E/W suites, where we use some of ECCC's functionality for the atmospheric forcing preparation and the mechanics of running the NEMO model, including managing restart files and outputs. We augment this baseline with functionality to prepare boundary forcing, extract data from the real-time data feed, generate the runoff forcing and prepare ice-concentration input files for the "ice-if" feature. Fallback strategies for missing data and persistence strategies for forecasts using gauge data are also implemented here.

2.6 ROBUSTNESS AND LIMITATIONS

For an operational model to be useful, it must be robust and not prone to failure. We have not assessed the models exhaustively in this regard; such testing is an ongoing process. However, some aspects have been explored:

- Where gauge data is needed as a model input, fallback mechanisms are implemented to mitigate missing or bad gauge data. Typically, this means we prepare a climatology for each gauge to stand in when the gauge data is unavailable. With these prepared ahead of time and with appropriate tooling to automate the switchover, the models can run despite missing gauge data and experience a graceful degradation through forcing with lower-quality data rather than a failure. Measuring the severity of the degradation under data-loss scenarios is reserved for future work.
- A long hindcast is conducted for model performance assessment. This long simulation demonstrates that the model is stable subject to a multi-year sample of weather/forcing conditions. In some cases, the hindcast period samples some extreme events, which helps bolster the case for model stability.
- Daily demonstration simulations (pseudo-analysis and four forecasts) have run for order one year on the General Purpose Science Cluster (GPSC) on a best-effort basis, to show that the automation suite can run the models routinely and reveal edge cases that can be fixed to improve robustness further. The purpose of running these pre-operational, best-effort simulations is to demonstrate the functionality/stability of the NEMO-based numerical model and the driving automation suite and identify issues that would impact operational deployment. The dominant source of issues experienced that impede on-schedule daily forecasts are (a) GPSC compute system downtime, both planned and unplanned, and (b) lack of availability of the forcing data from the larger-scale models that are nominally mirrored on schedule from ECCC's systems. These issues are deemed an expected consequence of using a research cluster and would be mitigated using an operational cluster.

The models will have some limitations:

- Intrinsic variability is expected in each model, and this has not been characterized.
- The lack of wetting and drying capability in NEMO 3.6 requires artificial bathymetry deepening in intertidal regions.

3 SAINT JOHN HARBOUR AND THE BAY OF FUNDY

3.1 REGIONAL OCEANOGRAPHY

The Bay of Fundy is a highly tidal bay lying between New Brunswick and Nova Scotia, connected to the northern extent of the Gulf of Maine. The geometry of the Gulf of Maine – Bay of Fundy system is near-resonant with the M2 tidal frequency (Garrett 1972), and so the tides in the region are large: in the upper Fundy region, the tidal range reaches 16 m, creating significant intertidal areas. Saint John, New Brunswick is about halfway up the Bay of Fundy, and the tidal range in the harbour is around 8 metres. The city of Saint John lies on the Saint John River, which is the largest river outflow in the Bay of Fundy. The river plume drives much of the near-surface circulation in the harbour, and as it propagates down the coast and dissipates, it provides a significant freshwater flux to the Gulf of Maine.

Non-tidal circulation at the entrance to the bay is driven mostly by flow from the Nova Scotia current that propagates along the South Shore of Nova Scotia and turns clockwise around Cape

Sable; smaller contributions from flow propagating through the Northeast Channel and from the Gulf of Maine circulation are also present (Aretxabaleta et al. 2008; 2009). The nontidal flow into the bay is predominantly on the eastern side of the entrance, and the predominant flow out of the bay is to the east of Grand Manan Island; a gyre circulation in the lower Fundy region is present throughout most of the year. The relatively narrow channel between Grand Manan and the mainland is predominantly an outflow, which is a source of freshwater for the coastal current in the Gulf of Maine.

Important locations around the Bay are shown in the top and middle rows of Figure 2, and the general near-surface circulation patterns described in the previous paragraph are shown in the bottom row. The top and bottom rows show the bathymetry of the outer domain (Fundy500) bathymetry, and the middle row shows the bathymetry of the inner domain (SJ100; in total on the left and zoomed in around the Saint John harbour on the right).

The city of Saint John lies on the Saint John River, which is an extensive river system flowing through the middle of New Brunswick. Both domain configurations include the Saint John River inland nearly up to Evandale, about 40 km inland from the Bay of Fundy. The river is generally between 5–20 m deep, with shallower regions more prevalent further inland. Kennebecasis Bay is a much deeper offshoot of the river, reaching nearly 75 m deep in some places. A weak tidal signal propagates to the extent of the river included in the domain, though the river is imposed as a fresh water open boundary and no salt is exported or imported from the system via the river open boundary. The tidal signal in the upper reaches of the domain (i.e., near Oak Point, water office gauge id 01AP003) is heavily modulated by the sill at Reversing Falls, which acts as a throttle modulating the interaction between the river and the ocean water. At low tide, the river flows from the land to the ocean, but at high tide, the tidal range on the downstream side of the sill is large enough that unless the river stage is high, the incoming water level from the harbour exceeds the water level signal from the river, and the river flows inland (see Figure 3). Tidal constituents show a sharp drop in amplitude and a phase lag between the harbour and just upstream of Reversing Falls (at gauge 01AP005), though these changes may not all occur solely at the sill (Horwitz, et al., 2021). As the tidal signal travels from the harbour gauge (CHS 00065) and the Reversing Falls gauge on the upstream side of the sill (01AP005), M2 drops nearly 3 m of amplitude and is lagged by about 30°. A double salt wedge forms due to Reversing Falls, and the significant presence of salt in the vicinity of Reversing Falls means that the (fresh) river boundary must be placed farther upstream. There is no salinity data available at Reversing Falls, and due to a systemic bias due to the change in density from the periodic presence of salt, it is not used for water level evaluation.

3.2 DOMAIN AND CONFIGURATION

The port of Saint John system (POPS-SJ) consists of two one-way nested domains downscaled from CIOPS-E and is an updated version of the system described in (Paquin, et al., 2020). The Bay of Fundy domain (Fundy500) refines the CIOPS-E grid by a factor of 5, and so has a nominal horizontal resolution of about 500 m. More precisely, the zonal resolution varies from 430-440 m, and the meridional resolution varies from 395–415 m. The Saint John Harbour domain (SJ100) refines Fundy500 by a factor of 5, and so has a nominal resolution of 100 m but more accurately a resolution of around 85 m. Since the grids are simply refined and not stretched or rotated, both Fundy500 and SJ100 are high-resolution ORCA grids; bathymetry for both domains is shown in Figure 2. Fundy500 covers the entire Bay of Fundy and ends at roughly a line of latitude spanning from the end of Digby Neck to the coast of Maine; Saint Mary's Bay is not included, and neither are any of the channels connecting Saint Mary's Bay to the Bay of Fundy. SJ100 has a smaller domain, extending from the Lepreau River to St. Martins and reaching about 40 km towards the middle of the bay.

Modelling the river system is critical to capturing the near-surface circulation patterns in Saint John harbour. Both Fundy500 and SJ100 include the Saint John River system from slightly upstream of Oak Point to the harbour, and Kennebecasis Bay is included. The river outflow is strongly modulated by the sill at Reversing Falls, which is about 5 km upstream from the head of the harbour.

Fundy500 and SJ100 share a vertical grid, which varies from dz = 1 m at the surface to dz = ~18 m at ~200 m depth. Fundy500 has 40 vertical levels, and SJ100 has 34. The bathymetry used for both Fundy500 and SJ100 is referenced to CGVD28 and was provided for this project by the Canadian Hydrographic Service (CHS). Both domains use the same gridded data as that used in (Paquin, et al., 2020), with very minor adjustments made to the open boundary regions to allow for smoother signal propagation from the parent model. The coastline data is the Canadian Continuous Vertical Datum Hydrographic Vertical Separation solution (commonly called CANEAST2015v1CL) (Robin, et al., 2016), and was provided by the CHS. As NEMO3.6 does not handle intertidal regions, a minimum water depth of 7 m is set throughout most of Fundy500 and all of SJ100 (including in the Saint John River system); a value of 9 m is used in the Upper Fundy area to account for the large tidal range. This unfortunately flattens out some topographic features near the coastline, and as such the model may not fully capture circulation that is funneled or constrained by shallow topographic features. To retain the resonant properties of the Bay, the extents of Minas Basin and Chignecto Bay are reduced to keep the total volume of water in the upper Fundy region consistent regardless of whether or not the intertidal zone is included in the numerical model. We also reduce the extent of Musquash Harbour, though this does not have a substantive impact on the tidal resonance of the system. The archipelago of islands and channels in Passamaquoddy Bay are also simplified in Fundy500 due to lack of resolution. Since the bay is open and the coastline does not feature fjords or narrow inlets, the HRDPS resolution of 2.5 km is sufficiently fine to resolve the near-shore atmospheric features in the region.

Important model parameters are provided to NEMO via a namelist file and are listed in Table 1. The advection scheme is total variation dissipation (TVD) with vertical substepping for stability, and the turbulence closure is generalized length scale scheme (GLS). Horizontal viscosity is handled by Laplacian viscosity with a Smagorinsky scheme. Nonlinear bottom friction is used with a quadratic formulation. Small patches of increased bottom drag are imposed around Reversing Falls for both Fundy500 and SJ100, and in Minas Passage for Fundy500; this helps maintain the stability of the model. Partial slip lateral boundary conditions are used, and the model conserves both energy and entropy.

3.3 INITIALIZATION

Initial conditions are generated from CIOPS-E for Fundy500, and from Fundy500 for SJ100. CIOPS-E does not have a river system, and initial values for Fundy500 in the river are coarsely interpolated from the open bay. Fundy500 is initialized on 2016-02-21 and to ensure that the river is sufficiently flushed of its initial conditions, it is run for over a month before SJ100 is initialized on 2016-04-01. SJ100 is allowed to spin up for 15 days, and the analysis period for both Fundy500 and SJ100 starts 2016-04-15. Both Fundy500 and SJ100 are then run until February 1, 2022, and forecasts are run daily from 2021-12-01 to 2022-01-31.

3.4 OPEN BOUNDARY CONDITIONS

The open boundary conditions (OBC) formulations are Flather (for SSH) and flow relaxation scheme (for the three dimensional T, S, U, and V). A combination of hourly data from CIOPS-E and WebTide constituents is used to force Fundy500. CIOPS-E data is detided via a set of (static) constituents that are calculated over a multi-year time span so as many constituents as

possible are included, however, neither SA nor SSA are included. Ten Webtide constituents from the Scotia-Fundy-Maine mesh (M2, N2, S2, O1, K1, L2, K2, 2N2, NU2, M4) are extracted at the Fundy500 boundary location and applied without any tuning (Dupont, Hannah, and Greenberg 2005). An additional four constituents (P1, λ2, T2, Q1) are calculated by inference from constituents K1, L2, S2, and O1 respectively, and are also applied. Gauge data from Yarmouth and Cutler-Farris are used to calculate the ratio between the inferred and reference constituent amplitudes and the difference of their phases, and the inferred constituents are calculated by scaling the source constituents by the average of the relationships at the two gauges (see Table 2). The domain is small enough that the relationships between constituent pairs at the two gauges are not drastically different. The four inferred constituents are each small (amplitude < 5 cm) but including these small constituents reduces the total water depth error in Saint John. Note that the M4 overtide is included in the OBC forcing even though it is predominantly locally generated; it is available in the WebTide dataset, and CIOPS-E includes it in its OBC forcing. M6 and other overtides are not explicitly included, and all other constituents that could have been inferred and applied have observed amplitudes smaller than 1 cm.

SJ100 is forced with hourly data from Fundy500. Before SJ100 is run, tidal constituents from two years of Fundy500 (April 2016–April 2018) are calculated using T_TIDE at the SJ100 open boundaries for use in both detiding and reimposing the tidal signal. (A span of two years is used to be able to include the T2 constituent.) The Fundy500 data is detided with these constituents, and then the same 14 constituents are reimposed. This is useful for two reasons: a) imposing the tides as constituents allows the tidal forcing to smoothly vary and avoids the signal clipping that results from forcing at discrete times, which would otherwise truncate the tidal signal by about 30 cm (~10%) for hourly data, and b) using isolated constituents allows constituents to be tuned carefully and individually, improving the results in the area of interest. M2 is tuned first by finely adjusting the bottom friction parameter, and no adjustment is made in the static constituent files. The other 13 constituents have adjustments to amplitude and / or phase applied, and these are detailed in Table 3. The adjustments are generally small, though the inferred constituents have proportionately larger adjustments made than the constituents supplied by WebTide, and phases generally required more tuning than amplitudes. The tide tuning was done iteratively, with an initial run done, then tides computed, and constituent files adjusted as informed by the discrepancy between the observed and modelled tides. This was repeated until adjustments did not substantively impact the results, and the net adjustment is reported here. The larger constituents were tuned before smaller constituents, and the inferred constituents were tuned last.

There is a mismatch in vertical datums between the port model system and the parent system (CIOPS-E), so we adjust the incoming lateral boundary conditions of Fundy500. The CIOPS-E domain extends beyond the definition of the CGVD28 surface, and the local offset between the CIOPS-E datum and CGVD28 is not known a priori. To determine the difference between the CGVD28 surface and the reference surface of CIOPS-E, we use long time series of observed tide gauge data from three water level stations in the region (Saint John 00065, Yarmouth 00365, and Halifax 00491). The observed mean sea level at each station is calculated, converted from chart datum to CGVD28 using known offsets provided by the CHS, and then compared to the modelled mean sea level from CIOPS-E at the three stations. The difference between the observed and CIOPS-E sea level is similar between all three stations: 0.687 m, 0.681 m, and 0.693 m for Saint John, Yarmouth, and Halifax respectively (Paquin, et al., 2020). To ensure that the system is consistent, the incoming CIOPS-E sea surface height is increased by 0.69 m. In Paquin et al 2020, the incoming ocean OBC were not adjusted, and instead the river water level (which is provided on the CGVD28 datum) was reduced by 0.69 m. This was done as there were concerns that adjusting the ocean OBC would disrupt the tidal dynamics of the system, however, we have since found that this is not the case. Using a consistent vertical

datum for all components of the system allows us to evaluate the model's total water depth, rather than just the tidal constituents and a de-meaned non-tidal component; this evaluation is critical for e-navigation purposes.

3.4.1 River OBC and NS_TIDE

The river boundary is applied just upstream of Oak Point using water level data from the Oak Point gauge (01AP003), a climatological temperature, and zero salinity. The boundary is imposed as a water level, and so the velocity is left to evolve freely from the slope of the surface. The flux from the river into the harbour is then controlled by width of the river channel: a wider channel allows more flux out of the river, with faster velocities due to a steeper slope in the surface. A temperature sensor is now available at Oak Point, but it was installed in 2021 and so for consistency the climatological river temperature was used for the entire hindcast.

The river OBC is implemented using data from gauge 01AP003, so as to capture the full variability of the physical river. This station is inland of the harbour, but there is still a detectable tidal signal in the data, superimposed on the longer time scale fluvial system; i.e., the Oak Point water level signal is a mixed fluvial-tidal signal. Additionally, using gauge data requires a) a fallback to patch gaps of missing data, and b) a forecasting procedure. Both of these requirements are met by using NS_TIDE, an extension of T_TIDE designed for use in rivers where the tides are not stationary (Matte, Jay, & Zaron, 2013). NS_TIDE uses a purely fluvial upstream station and a purely tidal downstream station (as well as any other locally relevant predictors) to decompose the signal at an intermediate station that has a mixed fluvial-marine signal. This decomposition is done by first generating a basis relating the fluvial, tidal, and mixed signals over a long time series. Care must be taken to avoid including any flooding events when generating the basis, as the relationship between the timing of signals at each of the gauges breaks down when water can flow over the banks as well as down the river. We use year-long time series from July 2019 to July 2020 to generate the lagged relationships between the gauges; the historic flood in the spring of 2019 had ended by late May. In addition to the fluvial upstream gauge (Fredericton, 01AK003) and the tidal downstream gauge (CHS gauge 00065), we include a third signal from the Reversing Falls gauge (01AP005), even though it itself is a highly mixed signal: without this, NS_TIDE has no information about how the tidal signal shifts over the sill. Omitting this gauge from the calculations nearly doubles the RMSE between the prediction and the observed signal at Oak Point, though this amounts from shifting to an RMSE of about 5 cm to one of about 10 cm. The gauges, their lags, and fallback data are all presented in Table 4. Since real-time data is not manually quality controlled, all gauge data is run through a threshold filter to eliminate bad data, and gaps in each gauge timeseries are patched using the method listed in Table 4.

Once this basis is generated, a prediction for the water level at Oak Point can be calculated by providing time series of upstream and downstream data, lagged appropriately. This data does not need to be restricted to the same window of time that was used to generate the basis, as the relationship between the gauges does not shift from year to year. The basis can be used to provide predictions in flood conditions, though the error is typically modestly larger than an ordinary time series of the same length; generating the basis is more sensitive to flood conditions than using the basis is. Figure 4 shows the generated basis as well as the observed gauge data, as well a two-day long prediction in spring freshet, low water, and fall freshet conditions. RMSE between the predictions and observations is typically less than 10 cm, reaching 10 cm only with short predictions in very high river stages.

3.5 FRESHWATER INPUT

The only freshwater input into the system is the Saint John River, which is implemented by a water level open boundary at Oak Point as described in Section 3.4.1. The Saint John River is the largest source of freshwater river outflow into the Bay of Fundy, and no other freshwater runoff is used in the model. Future model development can consider adding climatological freshwater input from other rivers that may be important to local circulation patterns or near-shore applications. These may include the Musquash, Lepreau, Annapolis, St. Croix, Shubenacadie, Salmon, and Petitcodiac rivers.

3.6 ICE MODEL

Ice is handled by NEMO's ice-if condition. The ice-if formulation checks if ocean water hits the local freezing point, and if it does, NEMO clamps the temperature and heat flux, and sets the ice concentration to a value specified by a provided file. The provided file used has a uniform concentration of zero.

3.7 MODELLING SYSTEM STABILITY

Stability is paramount for an operational model, and some basic assessment of the model's robustness was done. Firstly, the model ran for a six year hindcast without intervention with neither physical nor numerics-related crashes, and this six year interval included floods, hurricanes, and significant storms. The 2018 flood of the Saint John River was the largest flood on record, with the Oak Point water level reaching nearly 6 m; the typical freshet peak water level is around 3 m. As a crude stress test to ensure numerical stability, the model was run for the 2018 spring freshet / flood with water levels peaking at twice the observed height; the model remained stable and produced physically reasonable results. Realtime observations are used only in the river OBC calculation, and the fallback data for all gauges used is listed in Table 4. The atmospheric and ocean model data are supplied from operational systems, and so do not have fallbacks.

Limitations of the system can be loosely grouped into two categories: limitations due to the fundamental geometry of the system, and limitations due to the data used to force the system. The first category includes:

- NEMO3.6 does not account for intertidal areas, and so a minimum depth must be imposed throughout the system to maintain stability. This minimum depth may deepen shallow bathymetric features, which may reduce or adjust how currents are steered by the topography. Some features of the Saint John Harbour bathymetry are flattened by this, and so currents may be steered differently or insufficiently in the model as compared to observations. Results in the upper Fundy area may not be reliable due to the region's extensive intertidal flats.
- The connection between the river system and the harbour is quite narrow, especially at Reversing Falls. This is unavoidable for this configuration, since the highest resolution domain has dx ~85 m, and the river narrows to ~200 m between Reversing Falls and the harbour. The river system is included primarily to provide robust river outflow to harbour and the open bay, and circulation within the river system is not well validated.

The second category includes factors such as:

- Any biases or changes in variability in the parent forcing systems (CIOPS, HRDPS) will be inherited by the model.

- The river temperature is imposed as a climatology, which was derived from a variety of river temperature data in the area. This may affect the accuracy of the model's river plume temperature, which may have implications for the mixing of river and ocean water.
- Tides are imposed as static tidal constituents, but the specific values of M2 (and other) constituents is known to shift from year to year (Ku, Greenberg, Garrett, & Dobson, 1985). Any single set of constituents will then be of varying skill at representing the tidal water level in the model. The observed M2 amplitude fluctuates less than 10 cm over the course of the hindcast.

4 EVALUATION METRICS

The model performance is assessed through the analysis of a multi-year hindcast and a shorter set of forecasts, where the dates considered are constrained by available surface and boundary forcing from larger scale models. The hindcast evaluation uses a wide set of observations to analyze the model's representation of ocean conditions, including tidal analysis and model drift that would be difficult to assess on a short model run. Meanwhile, the forecast evaluation focuses on measuring the degradation of model skill as a function of forecast lead time for a smaller set of observations available during the forecast period.

Quality controlled data sources were preferred where possible. We performed additional quality control to some data as needed, including visual inspection, thresholding and automatic de-spiking to eliminate suspect data.

4.1 HINDCAST

Hindcast performance is assessed through comparison with available observational data. For each observation, we extract the corresponding virtual observation from the model. The error is defined as

$$\text{ERROR} = \mathcal{X}_m - \mathcal{X}_o,$$

where $\mathcal{X}_o$ and $\mathcal{X}_m$ are the observed and modelled values such that a positive/negative value indicates a model over/underestimate. For currents, $\mathcal{X}_o$ and $\mathcal{X}_m$ are taken as complex numbers with the real part representing the eastward and imaginary the northward components of velocity.

We use several scores, the bias, the centered root mean square error (hereafter CRMSE) and the root mean square error (hereafter RMSE),

$$\text{bias} = \frac{1}{N}\sum \text{ERROR} = \frac{1}{N}\sum \mathcal{X}_m - \frac{1}{N}\sum \mathcal{X}_o = \overline{\mathcal{X}_m} - \overline{\mathcal{X}_o},$$

$$\text{CRMSE} = \sqrt{\frac{1}{N}\sum (ERROR\text{-}bias)^2},$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum ERROR^2} = \sqrt{bias^2 + CRMSE^2},$$

and these measures retain the units of $\mathcal{X}$.

We also use the unitless gamma squared score,

$$\Upsilon^2 = \frac{CRMSE^2}{\sum(\mathcal{X}_o - \overline{\mathcal{X}_o})^2},$$

which is the ratio of error variance to observed variance, such that zero indicates perfect agreement between model and observation variance, and unity indicates error variance is as large as the signal variance. A value of unity or larger indicates no skill.

For scalar quantities, we use the unitless sample Pearson correlation coefficient score,

$$\text{PEARSONR} = \frac{\sum (x_m - \overline{x_m})(x_o - \overline{x_o})}{\sqrt{\sum (x_m - \overline{x_m})^2 \sum (x_o - \overline{x_o})^2}},$$

where zero indicates no correlation and unity indicates perfect correlation. For vector quantities (currents) we have the vector correlation coefficient (Kundu 1976; Röhrs and Christensen 2015),

$$\text{VECTORR} = \frac{\sum (x_m - \overline{x_m})^* (x_o - \overline{x_o})}{\sqrt{\sum (x_m - \overline{x_m})^2 \sum (x_o - \overline{x_o})^2}},$$

where the asterisk represents complex conjugation, which is also implied for squaring of complex numbers. This quantity is a complex number, where the magnitude measures the overall correlation and the angle is a measure of the average angle of the modelled current with respect to the observation. Here we compute the angle as positive clockwise to be consistent with the convention of representing the velocity direction as positive clockwise from North. Lastly, the vector correlation angle is only considered meaningful when the correlation magnitude is large (Kundu 1976).

4.1.1 Water level

Water level observations from the Canadian Hydrographic Service (CHS) are transformed from CHS Chart Datum to Canadian Geodetic Vertical Datum of 1928 (CGVD28) by subtracting a station-specific offset. CHS provides these offsets are based on measurements from a GNSS (Global Navigation Satellite System) occupation at each station. Pacific US stations are shifted from the US datum to CGVD28 via NOAA's VDatum and NRCan Vertical Datum Transformations online tools. Atlantic US stations were gathered with a mean low-low water datum and converted to CGVD28 using conversion surfaces provided by CHS. Water level observations from the Water Survey of Canada (WSC) are shifted to CGVD28 using offsets provided by WSC.

Tidal analysis is performed using T_TIDE on one-year segments at an hourly sampling interval. Where data is available at a higher frequency, it is interpolated/subsampled to hourly frequency. Observations missing more than 10% of the data points are not detided. The tidal analysis enables the decomposition of the total water level into tidal and non-tidal parts as

$$H_{total} = H_{tidal} + H_{nontidal}.$$

The annual (SA) and semi-annual (SSA) constituents are excluded from the tidal analysis; these constituents are subsumed into the non-tidal component. For the present purpose this is adequate as the same process is applied to both model and observations, and the extent to which the models do not reproduce SA or SSA is captured in the non-tidal error scores. Constituents with signal-to-noise ratios below 2 are also subsumed into the non-tidal part. Water level bias is included in the non-tidal component. Scores are reported for the total, tidal and non-tidal components. Owing to a negligible cross-correlation between tidal and non-tidal water levels, to a good approximation, we have

$$\text{CRMSE}^2_{total\ wl} \cong \text{CRMSE}^2_{tidal\ wl} + \text{CRMSE}^2_{nontidal\ wl},$$

which shows how these two errors contribute to the total. Amplitude and phase errors are reported for the significant constituents, as is the tidal error,

$$\text{TIDAL ERROR} = \left[\frac{1}{2}({h_o}^2 + {h_m}^2) - h_o h_m \cos(\phi_o - \phi_m)\right]^{\frac{1}{2}},$$

where h_o and ϕ_o are the observed (subscript *o*) amplitude and Greenwich phase lag for a given tidal constituent, while h_m and ϕ_m are the modeled (subscript *m*) amplitude and phase (Cummins and Oey 1997).

4.1.1.1 Storm surge evaluation

To evaluate the port models' performance during storms, a handful of stormy periods are selected for each port. The cyclone database of (Zhang, Perrie, and Long 2019), which provides global storm tracks from 1958 to 2021 (Figure 5 shows a zoom of North America), guides the selection. The database was queried for storms passing within a few hundred kilometres of each port model domain, and it provides a comprehensive list of storm events during the hindcast period. However, this list is quite extensive for some ports (particularly Canso and Saint John), so longer lists are narrowed down by considering minimum central pressure, examining water level gauges, and local media reports of storm impacts. The evaluation follows the residual water level hindcast evaluation in terms of performance scores and plots but is limited to the storm periods.

Storm surge detiding is done using a 40-day window around the storm's peak, i.e., 20 days before to 20 days after. This differs from the typical water level analysis, which is done in yearly spans and as such can include more constituents. However, using a shorter window does a better job removing the tides during each storm, particularly in areas with non-stationary tides.

We note that storm surges may interact non-linearly with tides to impact water levels, but we do not investigate this phenomenon here.

4.1.2 Water velocity

Velocity data is available from ADCPs and some current metres in one of four configurations: mounted on a mooring, buoy mounted (i.e., floating), in a float towed by a ship, or horizontally mounted on a shore structure (HADCP). Only horizontal (east-west / north-south, u/v) velocities are considered here. The horizontal velocity's u/v and speed/direction decompositions are both considered, and directions are calculated as positive clockwise from north.

Model data is extracted to match the observed data's time span and spatial location. Time series longer than 29 days and with less than 10% of data points missing are detided using T_TIDE; hourly data is used for detiding, and where more frequent data is available, it is first down sampled. As with water level, the SA and SSA constituents are excluded from fitting, a Rayleigh value of 2 is used, and the observations and model data are processed the same way.

Currents are noisier, tend to have more missing data, are less stationary, and are impacted by more non-linear processes than water level observations, and these factors make the detiding process less robust. Owing to nonlinear processes such as bottom friction, some tidal (kinetic) energy will spread into adjacent frequencies, so fitting to tidal harmonics does not capture all variability induced by the tides, and some near-tidal variability remains in the residual velocities. We do not apply low-pass or band-stop filters to the residual to reduce the noise. While doing so would provide a more completely de-tided residual time series, some of the total signal would be unaccounted for by either the tidal or non-tidal evaluation. Given these caveats, the tidal component may be more accurately described as "the component of the currents that T_TIDE is able to fit." Currents are evaluated using similar metrics to water level, using complex

formulations. Tidal constituents are evaluated using the ellipse error (Cummins and Thupaki 2018)

$$D_u = \left[\frac{1}{2}(A_o^2 + B_o^2 + A_m^2 + B_m^2) - \cos(g_o - g_m)\cos(\theta_o - \theta_m)(A_o A_m + B_o B_m) - \sin(g_o - g_m)\sin(\theta_o - \theta_m)\,(A_o B_m + A_m B_o)\right]^{1/2},$$

where A and B are semi-major and semi-minor axes, respectively, the subscripts *o* and *m* correspond to observed and modelled, g is phase, and θ is angle of inclination. For the non-tidal and total time series, complex formulations of the metrics listed at the beginning of the section are used, including bias, CRMSE, γ^2, and vector correlation.

We note that for single-location instruments, in particular moored current meters and moored and horizontally mounted ADCPs, we expect the scoring to be sensitive to the details of the model run under evaluation. Small errors in bathymetry can adjust the location of deterministic features (eddies, jets, meanders, etc.) and this can lead to large errors in the scoring. Meanwhile, chaotic internal variability can also affect the location of such features. An ensemble of model runs could help mitigate the latter but is beyond the scope of the current effort.

4.1.2.1 Horizontal ADCPs

HADCP data is decomposed into along- and cross-channel components, and evaluation is done primarily with mid-beam data to avoid edge effects. When a long enough time series is available, a tidal analysis is done as with moored instruments. Time series of the first week of data are plotted, regardless of how long the total time series is, to show the daily variability in the signal and how well the model captures it. Scatter plots are used to show the distribution of speed and direction of the total velocity, and the semi-major axis for the tidal ellipses are compared when feasible.

4.1.2.2 Current meters and moored ADCPs

Current meters report velocity at a single depth, so those velocities are evaluated at instrument-specific depths. For ADCPs on moorings or buoys, a standard set of depths at which to evaluate the velocities is selected on a per-port basis, taking into account local bathymetry and data availability; for brevity, statistics may be reported only at some depths for each instrument considered. These levels are defined relative to either a nominal surface (i.e., one with no sea-surface height variation) or the time varying sea surface. For regions with ADCPs in relatively deep water without a large tidal range, using the nominal surface is sufficient. For shallow regions and/or those with large tidal ranges, the evaluation is done relative to the time-varying sea surface, using the observed total water depth to process the observations and modelled sea surface to process the modelled currents.

The observations are preprocessed to remove any spurious data points: values larger than 10 m/s are discarded and the data points are resampled if needed to be evenly spaced.

Tidal analysis is done on all timeseries of at least 29 days at depths with less than 10% of the data missing. Tidal ellipses are plotted for the largest tidal constituents, including depth profiles of tidal ellipses for the largest constituents. Time series and histograms are plotted for both the non-tidal and total velocities. Analysis of total velocities only is done for shorter time series or time series with missing data. We note that if data is missing at consistent phases of the tide, then the analyzed results may be aliased, and so time series with substantive regular gaps are not analyzed or presented.

4.1.2.3 ADCP transects

ADCP transects are too short for tidal analysis, so only the total velocities are considered. An along-transect / cross-transect decomposition is used, with a constant angle used for decomposition for the entire transect; this is sufficient as the transects are typically short and straight. Along-transect plots of velocities at depth are used for evaluation. Scatter plots of bias vs CRMSE are also plotted, with the statistics calculated at each physical point and integrated over the transect.

4.1.3 Water properties

Conductivity-Temperature-Depth (CTD) profiles are grouped into manually defined subregions based on the geography of each port domain. This enables an aggregate assessment over areas, including bias and CRMSE as a function of depth for each area. Model results are taken from the nearest point to the data location and nearest to the observation time. Vertical profiles are interpolated to the model z-levels to bring all data to consistent depths.

Sea surface temperature buoy and moored CTD measurements are evaluated using the bias, CRMSE, χ^2 and Pearson's *r* scores over the evaluation period. Model values are linearly interpolated to observation locations in the horizontal and vertical dimensions. The observed and modelled time series are interpolated to the largest common data interval.

Ferry thermosalinographs are vessel-mounted temperature and conductivity sensors. They provide a measure of near-surface water temperature and salinity by collecting data through one of the vessel's seawater circuits. Model results for comparison with these data are taken from the point nearest to the observed locations both horizontally and vertically and then linearly interpolated to the observation times. Hovmöller plots of observations, model results and differences, plus basic statistics, are used to assess the model performances relating to near-surface water.

4.1.4 Drift

The models' performance in drift trajectory prediction is assessed by comparing the observed tracks of ocean surface drifters with analogous trajectories modelled using the surface currents output by the port models, in combination with wind forcing from the National HRDPS atmospheric forecast. This comparison is done using the drift evaluation tool developed as part of the DPNM sub-initiative, in the OpenDrift configuration (Soontiens and Holden 2024).

Windage on the surface drifters is applied by computing the wind drag coefficient based on the drifters' drag area ratio (Niiler et al. 1995; Daniel et al. 2002; Röhrs et al. 2012; Hourston 2021; Blanken et al. 2021). This coefficient parameterizes the effect of direct wind drag on the parts of the drifter exposed above the water surface and varies with drifter geometry. To account for wave-induced Stokes' drift, an additional 1% of the wind speed is added (Sutherland et al. 2020), and this sum is applied as the windage in the trajectory prediction.

This method of applying windage assumes that currents over the draft of the drifter are known exactly, as are winds directly at the ocean surface. However, in reality, neither of these assumptions is true, since:

1. The representation of surface currents in NEMO is limited by stability restrictions on near-surface vertical resolution and model uncertainty in general; and
2. Wind speed from the HRDPS model is also subject to model uncertainty and given at 10 m where winds are usually ~30% larger than at 1 m but could be as much as five times larger depending on wind speed and atmospheric stability (Smith 1988).

The representation of Stokes' drift as 1% of the wind speed represents a further assumption, as this value was derived by comparing tracks from various types of surface drifters to currents from the Regional Ice-Ocean Prediction System (RIOPS) ocean model and Canadian Arctic Prediction System (CAPS) atmospheric model (Sutherland et al. 2020). As noted in the discussion section of (Sutherland et al. 2020) and references therein, this value can vary widely depending on the combination of ocean and atmospheric forcing, which implies that model uncertainty contributes significantly to the appropriate value here. The authors also note that explicitly including Stokes drift based on a wave prediction system is preferable over parameterization based on wind velocity.

The windage term used here could be optimized by deriving it for each ocean/atmospheric model combination using the procedure in (Sutherland et al. 2020). However, this is beyond the scope of this report: the focus is on improvements to the current forcing for drift simulations without evaluating the suitability of wind predictions in the port model domains for drift prediction or commenting on the potential utility of a port-scale wave prediction system. The windage parameterization used here is merely intended to provide a consistent, deterministic linkage between modelled currents, winds, and the motion of various drifter types. As considerable uncertainty is associated with this windage term, evaluation of surface currents against observations from ADCPs and current meters is deemed the primary determinant of model suitability for drift prediction, rather than analysis of observed and modelled drifter tracks.

Periods where drifters were active in the model domain are identified, and modelled trajectories are started every hour along the drift tracks. The benefits of starting drift tracks in this manner are to reduce the sensitivity to initial conditions and increase the number of tracks available. However, it means that some drift tracks are not independent and thus the errors may be correlated. Modelled trajectories were computed for a user-specified period of 24 hours or more, where possible. However, in some regions, the majority of the observed drifter tracks were less than 24 hours long, and here a shorter modelled trajectory length was chosen.

Observed drifter tracks were truncated to areas covered by the 'wet' cells of the port model domains to avoid launching virtual drifters in 'dry' parts of the domain where observed drifters are near the shoreline, which may not be precisely resolved. In addition, to facilitate interpolation of starting locations for virtual drifters, observed trajectories were split where time gaps between position records exceeded two hours. In the remaining portions, position records were interpolated to a consistent time interval ranging from five minutes to one hour.

For each model – observation pair of trajectories, two statistics are computed to assess the model performance. First is the separation distance, D, which is given by,

$$D(t) = |\dot{x_o}(t) - \dot{x_m}(t)|$$

Here $\dot{x_o}$ and $\dot{x_m}$ are the positions of the observed and modelled drifter, and || denotes the magnitude of the vector difference, i.e., the distance, between them.

Second is the instantaneous skill score, S, following (Molcard et al. 2009), which is given by

$$S(t) \;=\; \max\left(0, 1 - \frac{D(t)}{d_o(t)}\right)$$

Here $d_o(t)$ is the displacement of the observed drifter from the starting point of the pair. The rationale for the normalization by $d_o(t)$ is to increase the skill assigned to a trajectory prediction as the trajectory length increases, even if the separation distance remains constant. A separation distance of, for example, 500 m represents a less grievous error in a trajectory that is 10 km long than in one that is 500 m long. A value of S=1 indicates a perfect prediction.

4.2 FORECAST

The forecast evaluation involves running a set of forecasts (here 48 hours long, each starting at 00Z) and evaluating the performance over the independent forecasts as a function of lead time. Forecast evaluation was performed for a set of order 60 consecutive forecasts by comparing the model values with tide gauge, sea surface temperature and horizontal ADCP records. The set of forecasts were taken from winter 2021/22 for logistical reasons. Forecast performance was evaluated as the discrepancy (bias and CRMSE) with observed values as a function of forecast lead time. The error growth curves represent the discrepancy averaged over the set of evaluated forecasts, and we include 95% confidence intervals computed with a bootstrap method.

To detide the forecast and the corresponding hindcast water level series we subtract the tidal signal precalculated based on the hindcast covering the forecast evaluation period. The tidal signal is obtained with a T_TIDE fit with a Rayleigh number as low as 0.1 (overfitting) to ensure maximal energy removal at tidal frequencies. Such strong suppression of tidal energy was implemented to clear the error growth curves of any tidal residual, which otherwise would dominate the curve.

5 HINDCAST EVALUATION RESULTS

To evaluate our two-level port model system, we compare our model results against CIOPS-E. We have hourly 3D data from CIOPS-E over the entire Fundy500 domain, and so we can make meaningful comparisons for all types of observations, not just water level and surface quantities. The hindcast evaluation window starts 2016-04-15 and ends on 2022-01-01, though the hindcast run is extended until 2022-02-01 to facilitate the forecast evaluation. The hindcast is divided by calendar year, so the year2017 evaluation period runs from Jan 1, 2017, to Jan 1, 2018. (Note that year2016 spans only ~8 months.) The entire hindcast period is labelled years16-21.

Observations used to evaluate the model include water level from tide gauges (both Canadian and American), temperature and salinity time series from buoys and moored CTDs, velocity from moored and towed ADCPs and floating HADCPs, depth profiles of temperature and salinity from CTD casts, and drifters.

5.1 WATER LEVEL

Circulation and water level in the Bay of Fundy are dominated by the large tides, which result in a tidal range of nearly 16 m in the uppermost reaches of the bay. Saint John is about halfway up the bay, and the tidal range is around 8 m; the semidiurnal constituents account for about 90% of the tidal range in Saint John. To be a useful model for operational purposes, the water level must be well represented by the model, though due to the lack of wetting and drying in NEMO 3.6, we do not expect to have reliable results in regions with extensive tidal flats (i.e., the upper Fundy region). There are some intertidal zones around Saint John which have been accommodated by setting a minimum bathymetric depth of 7 m in the simulations, but no validation is carried out in intertidal areas.

The port of Saint John is the predominant focus of this model and has a single real-time water level gauge (Saint John, 00065) located on a wharf with no intertidal areas nearby. Other real-time gauges in the Fundy500 domain are found in Passamaquoddy Bay (Eastport, NOAA gauge 8410140) and on the northern coast of Maine (Cutler-Farris, NOAA gauge 8411060); only the harbour gauge 00065 is present in SJ100. (See Figure 6 for the tide gauge locations.) A gauge at Parker's Cove (00320) on the Nova Scotia side of the Bay of Fundy was in operation for a

few years, with observations ending in 2017. Vertical datum conversions from chart datum or mean low-low water to CGVD28 are listed in Table 5.

As noted in Section 3.4.1, CIOPS-E has a different vertical datum than CGVD28, and long term mean SSH values at Saint John, Yarmouth, and Halifax indicate that the local difference between the CIOPS-E and CGVD28 water levels is about 0.69 m. All CIOPS-E water levels are raised by 0.69 m before comparison with Fundy500 or SJ100 results, and this allows for easier visual comparison between the model solutions.

5.1.1 Mean sea surface height

A long term mean sea surface height field is required by the CHS for their e-navigation products, and these fields for Fundy500 and SJ100 are shown Figure 7 and Figure 8. Throughout most of the domain the mean SSH is around 0.25 m, while the river system has values closer to 1.5 m. This is as expected: the river values are broadly consistent with the observed water level at Oak Point, and the sill at Reversing Falls both throttles the flow of river water into the harbour and significantly reduces the amplitude of the tidal signal propagating inland. The mean water level is higher in the Saint John harbour than through most of the rest of the domain, likely due to the river outflow and the retention of water due to the sill at Reversing Falls. Local minima of mean SSH are found off Grand Manan and in Minas Passage.

5.1.2 Tidal water level

M2 is by far the dominant constituent in the Bay of Fundy, though N2 and S2 are also significant. Tables of all forced constituents with amplitudes > 5 cm are found in Table 6 to Table 16 sorted broadly in order of amplitude at Saint John; constituents M4, T2, and Q1 are small enough to be omitted. Values from 2017 are shown so that Parker's Cove is included in the tables; the difference in parameter values between 2017 and any other given year is generally small. All tidal constituents from Fundy500 (except M2) are manually tuned before being used to force SJ100, which allows us to generate a precise tidal solution at Saint John. M2 tidal parameters for 2017 for all four stations are plotted in Figure 9. Given the careful work done to tune the tides at Saint John, it is unsurprising that SJ100 outperforms CIOPS-E there, but Fundy500 outperforms CIOPS-E as well. The amplitude of M2 is significantly improved by both Fundy500 and SJ100 at all stations, and the phases are comparable or improved as well.

Saint John, Eastport, and Cutler Farris all lie on the western shore of the Bay of Fundy; the only water level observations available on the eastern shore are at Parker's Cove and end in mid 2017. Even though the Parker's Cove timeseries is limited, comparing tidal constituents on both sides of the bay provides some insight into whether the cross-bay structure of the tidal solution is accurate. The Fundy500 domain provides comparable results to CIOPS-E: the amplitude of M2 is improved from CIOPS-E but is not as accurate as at Saint John. The phase errors for the largest constituents are relatively small, suggesting that the structure of the isophase lines across the bay are broadly correct.

Using static tidal constituents to force the tidal flow is well founded physically, as the tides from the parent CIOPS-E model are not sufficiently skillful to be used to force the near-shore model. However, because the tides in the Bay of Fundy are large due to the geometry of the system being resonant with the M2 frequency, the nodal modulation of the astronomical tide is reduced due to the physics of quadratic bottom friction. Our analysis tools account for the astronomical nodal modulation but cannot account for a locally specific modulation. Ku et al suggest that although the astronomical modulation is +/- 3.7 %, in the Bay of Fundy the nodal modulation is only +/- 2.4%, i.e., there is a 1.3% difference (Ku, Greenberg, Garrett, & Dobson, 1985). M2 at Saint John has an amplitude of around 3.05–3.10 m, and so we anticipate that our calculated

tidal amplitudes will vary by ~3–4 cm between years. Table 17 shows the M2 tidal parameters at Saint John for each of the six years examined, as well as the average. While the phase does not fluctuate more than one degree, the observed M2 amplitude ranges from 3.02 m to 3.07 m, which is broadly consistent with the difference we expect from our tidal analysis not accounting for the reduced nodal modulation. Our port model results show the closest match to M2 in the early part of the hindcast, i.e., the interval over which the constituents were calculated before being imposed for the full six years of the hindcast. In later years, particularly when the observed M2 amplitude is low, the difference between the modelled and observed M2 amplitude is larger. There is no clear way to avoid this problem when using static constituent files.

Historical tidal constituents are also available at over 50 stations around the Bay of Fundy; these are considered to be static and are provided by the CHS. (We use a subset of 35 stations distributed around the bay for more readable plots.) A map of the real-time gauges and constituent stations used in this evaluation is shown in Figure 6; constituent stations along the Saint John River are omitted in the evaluation, as well as stations that are not wet in Fundy500. Note that upper Fundy stations are included, to provide some estimate of model skill outside of where we expect our models to perform well. Amplitude and phase for M2 are plotted for these historical stations in Figure 10. Fundy500 and SJ100 consistently outperform CIOPS-E throughout the lower bay, from stations Cutler (00004) to St Martins (000129), and from Margaretsville (00315) to West Narrows (00339). Results near Saint John are particularly skillful in SJ100, with amplitudes within 5 cm of the static constituent. In the upper Fundy area, we underestimate both the phase and the amplitude of M2; the phase is off by only about 10° and the difference amplitude is at most 50 cm, and often only 10–20 cm. We did not expect to be skillful in this area and anticipate improvement when we move to NEMO4 with wetting and drying.

5.1.3 Overall scores

Statistics on the total and tidal water level for 2018 for CIOPS-E, Fundy500, and SJ100 are presented in Table 18 and Table 19. 2018 was chosen as a representative year because of the historic flood in the spring, which saw water levels at Oak Point peak at almost 6 meters, whereas a typical spring freshet peaks at around 3 meters. A second flood in the spring of 2019 was not quite as extreme, but severe flooding is becoming a more common issue in the area, and the river outflow is a critical part of the circulation in the harbour. Both Fundy500 and SJ100 use gauge data to provide river stage, while CIOPS-E uses a climatology, and cannot so account for interannual variability.

Plots of the statistics calculated over each of the six years covered by this hindcast are shown in Figure 11, and in general the statistics do not degrade significantly over the span of the hindcast for any of the three models shown, i.e., the scattered dots stay clustered for all stations. We find that Fundy500 and SJ100 both robustly outperform CIOPS-E in water level statistics, and SJ100 performs slightly better than Fundy500. The total water level CRMSE in SJ100 (i.e., the highest performing domain) is around 10 cm for all years, and this is driven equally by tidal and nontidal water level, which both have values between 5 and 10 cm for all years.

5.1.4 Non-tidal water level

Time series of non-tidal water level for Saint John, Parker’s Cove, Eastport, and Cutler Farris are plotted in Figure 12, Figure 13, Figure 14, and Figure 15 respectively. The time series for 2018 is shown for Saint John, to better see higher frequency variation, but the entire available time series are shown for the other three gauges. CIOPS-E data has been increased by 0.69 m so it can be easily compared to the other time series. The high frequency variation makes details of the time series difficult to see, but the histogram of values is helpful. SJ100 and

Fundy500 show similar results for Saint John and match the observations well, and Fundy500 also matches observations at Eastport and Cutler Farris. However, it does not perform as well at Parker's Cove: the observed histogram is sharper and narrower than the Fundy500 histogram. This suggests that the circulation along the Nova Scotia coast may benefit from some attention, and results may improve if outflow from the Annapolis River is included. Statistics for non-tidal water level for 2018 are listed in Table 20; statistics from other years are similar.

Frequency spectra for both the total and non-tidal water level for at Saint John and Parker's Cove for 2016 are shown in Figure 16 and Figure 17 respectively. The spectral analysis is sensitive to gaps in oscillatory signals, as it linearly interpolates over gaps in data. For the Saint John gauge, the 2016 analysis period has the least missing data of any of the years considered, and so has the least likelihood of including spurious energy generated by linear interpolation over substantive gaps in data; Parker's Cove is included for comparison. The total water level spectrum shows two things that have not yet been established: a) the overall magnitude and small features in the low frequency band are well captured by the port model, and b) the observations have a higher baseline of high-frequency energy (> ~3 cpd) that any of the models. These are both accurate for all stations in the domain (Eastport and Cutler Farris are not shown) but Saint John in particular has relatively more energy in the very high frequency band than any of the other stations; as such the discrepancy between model and observations at high frequency is largest for Saint John. This discrepancy may be a numerical artifact of the spectral processing, as the non-tidal spectra show a smaller discrepancy at high frequencies, and the non-tidal time series is not oscillatory. The nontidal spectra in Figure 16 underscore that we are matching the observed low frequency signal at Saint John better with SJ100 than either Fundy500 or CIOPS-E, particularly in the 0.1–1 cpd band; Parker's Cove sees a smaller improvement. However, all spectra at both stations lie within the 95% confidence interval so this cannot be considered a robust result, but the results are nonetheless encouraging.

5.1.5 Storm surge water level

The Bay of Fundy is frequently stormy, and over the course of the hindcast period several hurricanes, post-tropical storms, and nor'easters passed through the region. A selection of 11 storms spanning the hindcast years, seasons, and about 40 hPa of minimum central pressures (963–1008 hPa) were chosen to evaluate how well the hindcast captures storm surge. Storms are labelled by the month and year in which they occurred, or by name if it was a named storm or its remnants. The peak of the storm was taken to be the timestamp of the lowest central pressure, and the water level signal was detided using a 40 day window (i.e., 20 days before the storm peak and 20 days after). The surge signal is taken as the total water level minus the tidal signal. Plots showing a window of 4 days around the peak of each storm are shown for Saint John in Figure 18.

Our models reliably capture the timing of the surge signal, with both Fundy500 and SJ100 showing very similar time series. SJ100 tends to capture the peak water level slightly more reliably than Fundy500 and have a smaller overall bias, though the difference between the two configurations is small. Hurricanes are well modelled by both systems, and the signal after the peak of the storm has passed is better captured in Fundy500 and SJ100 than it is in CIOPS-E. The visual differences between the three models are small, but it is clear that the port model configurations perform at least as skillfully as CIOPS-E and capturing the timing and variability of surge events.

Statistics for the three long term water level stations are listed in Table 21; Parker's Cove is included but only one of the 11 storms considered occurred while the gauge was in operation. The port models match or out-perform CIOPS-E statistically at all stations, and SJ100

consistently outperforms Fundy500. This indicates that the higher resolution models improve on surge modelling over the existing system, even outside of the main region of interest.

5.2 WATER VELOCITIES

Although we expect the port model to provide improved results for water properties and water level, we expect the largest relative improvement to be in water currents. The river plume from the Saint John River heavily influences the near-surface circulation in the harbour, and we expect that higher resolution will improve the representation of the extent and structure of the plume. Visual observations from field work and satellite photos show that the edge of the river plume is a very sharp front, and sharp fronts persist in the circulation as the river plume is pushed back and forth by the tidal circulation. Figure 19 shows the surface speed at 2020-06-01, 00:00 UTC for CIOPS-E, Fundy500, and SJ100, and the effect of improved resolution is clear. Fronts in SJ100 are sharp, while they are fuzzy in Fundy500 and may not be visible in CIOPS-E. More small scale features are visible in SJ100, and patterns are more complex.

Velocity observations are primarily provided by ADCP deployments around the Saint John area in 2016, 2017 and 2019. There is also a downward looking ADCP mounted on the Smart Atlantic buoy that, aside from some interruptions, provides currents for the entire hindcast period. A handful of ADCP transects were conducted in the Saint John harbour in April 2019, and two horizontal ADCPs (HADCPs) were installed on floating rails mounted on wharf structures at the junction of the river and the harbour.

The Saint John harbour is relatively shallow, and the total water depth can vary by a factor of two over the course of a tidal cycle due to the large tidal range. All velocity timeseries are considered at a set depth below the variable surface, rather than being extracted relative to a fixed height above the bed. This avoids large gaps or aliasing in the time series by consistently missing data at low tide (see Figure 20 for a schematic). As such, often only relatively thin sections of water that are consistently analyzable, as the moored ADCPs are generally a few metres above the seabed in 10–15 m deep water.

5.2.1 Horizontal ADCPs

Two HADCPs were installed at the junction of the river and the harbour, one on each bank (see Figure 21). The one on the east side was installed at the old sugar refinery and the one on the west side was installed on Wharf #11. The HADCPs were set up to capture near-surface currents, and so to ensure that surface data would be captured throughout the tidal cycle, the HADCPs were installed on floating carriages that raise and lower on a pair of rails as the tide goes in and out. HADCPs were in the water in 2019 and 2021, though the 2021 deployment had some problems and yielded limited useable data.

The HADCPs are installed at a narrow point at the head of the harbour, and only SJ100 has sufficient resolution to give useful model results to compare to observations. Fundy500 uses a single grid-point wide channel to represent the river from just upstream of Reversing Falls to the harbour (and this grid point is slightly north of the actual river), and CIOPS-E does not include the river at all. All plots for this section show only SJ100, as the other two model domains have land in the relevant locations. Both the observed and modelled data are decomposed into along-shore and cross-shore velocity, and we focus on the along shore velocity. The heading of the east side instrument is 257° relative to true north, and the heading of the west side instrument is 50°.

Tidal ellipses from the first deployment of the west and east instruments are shown in Figure 22 and Figure 23 respectively; results from subsequent deployments are nearly identical. As expected, M2 is by far the dominant constituent, and so only the large semidiurnal constituents

are plotted. Both instruments have similar differences in inclination of the tidal ellipses, suggesting that the angle at which the river joins the harbour may not be precisely accurate; this may require higher resolution to resolve. The magnitude of the eastern ellipse matches observations well and has a hint of a width to the ellipse. The western ellipse has too long a semi-major axis and lacks the very slight width seen in the observations. Given the SJ100 domain's coarse resolution relative to the width of the channel, these results are acceptable.

Time series of along-shore total and residual velocities are shown in Figure 24 and Figure 25 for the western instrument, and Figure 26 and Figure 27 for the eastern side. The timing of the tidal cycle is well captured in all four time series, with the sharp rises and drops in along-shore velocity occurring simultaneously in the observations and model results. However, the model results are much smoother when the velocities peak than the observations are, particularly on the western side of the river where the model tends to overshoot the observations. The observations have an interesting structure: on ebb tide (i.e., when the along-shore velocities are most negative), there is almost a rebound, with the observations hitting a peak and then relaxing slightly (becoming less negative) before low tide. This is more prominent in the December deployment rather than the July deployment, and the December residual velocity shows much stronger periodicity than the July deployment. Sections of two deployments are not enough to draw firm conclusions about the impact of seasonality in the along-shore currents, but this suggests that there may be different features that arise depending on the river stage.

The eastern instrument is similar in that the model has a smoother signal than observations at flood and ebb tide, but the observations are flat and noisy at flood and ebb, rather than the inverted U shape seen in the western instrument. The model is generally less negative than the observations, indicating that we're underestimating the alongshore current on the eastern side of the channel particularly on flood tide. This is more prominent in October than it is in June, again hinting at some seasonality. While the error in the western instrument oscillates around 0, the eastern observations are consistently underestimated by the model.

Although we focus on the along-shore component of the velocity, the observations suggest that the cross-shore velocity may be interesting: the observed tidal ellipses do have small semi-minor axes, and there are high frequency oscillations visible especially at the edges of the peaks of the along shore velocities. An ADCP transect taken across the river roughly between the two HADCP locations shows cross-channel variability (see Section 5.2.4 and Figure 54) that is not well captured in SJ100, as there is not sufficient resolution to properly model much cross-river variability. A deeper delve into the observed HADCP data may be informative for future model development in the Saint John region, particularly if higher resolution nests are considered.

5.2.2 Current meters

There were no single-depth current metres present in the domain to use for model evaluation.

5.2.3 Moored ADCPs

There are about 25 ADCP deployments in and around the Saint John harbour during the course of the hindcast (see Figure 28). Most of these are short deployments of bottom-fixed upward-looking instruments of a few months, though there is also a downward-looking ADCP on the Smart Buoy which provides a multi-year time series; some locations are sampled multiple times. Most of the deployments available are near the harbour in shallow water, with a total water depth that may fluctuate by a factor of two over a tidal cycle. As such, the observations must be handled in a way that robustly accounts for the fluctuating surface. For this evaluation, all depths are measured relative to the position of the fluctuating sea surface, and not relative to a

fixed zero-SSH surface. If this were not done, the shallow water and large tides would cause periodic gaps in the data at low tide, which would bias the time series analysis in a non-physical way.

As with water level, M2 is by far the dominant constituent in the tidal currents. N2 and S2 are the next largest constituents but have semi-major axes about an order of magnitude smaller than that of M2 for all moorings considered; diurnal constituents are even smaller. SJ100 substantially outperforms both Fundy500 and CIOPS-E: Fundy500 has a crude river geometry that should be revisited in future versions, and CIOPS-E lacks both the harbour geometry and the distinct river plume that influences much of the near-surface circulation. Most plots in this section include CIOPS-E, but for visual clarity some timeseries plots may not. The circulation is evaluated at several depth-below-surface values, to maximally capture the baroclinic structure of the tidal ellipses.

5.2.3.1 Tidal constituents

The circulation in the domain is heavily tidal, and as such the port model's skill at matching the tidal ellipses is critical to reproducing the circulation patterns. Vertical profiles of M2 tidal ellipses for a subset of ADCP moorings are shown in Figure 29 through Figure 35. These moorings are located throughout the harbour (IH, OH, PI), at the Smart Buoy near the pilot boarding station, at Canaport (598), off Cape Spencer (607), and in Mace's Bay (612).

CIOPS-E severely underestimates the amplitude of the semimajor axis in the harbour and overestimates it at Canaport and Cape Spencer. CIOPS-E does not consistently capture the angle of the ellipses either and the ellipses do not vary significantly with depth. Fundy500 improves on CIOPS-E, particularly in the areas just outside of the harbour (for example, at the Smart Buoy or Canaport). The width of the ellipses and the angle of inclination are generally improved, and more baroclinicity is captured. However, SJ100 consistently outperforms both CIOPS-E and Fundy500 inside and outside the harbour: the baroclinicity is better captured, the amplitude and width of the ellipses are more accurate, and the angle of the ellipses matches the observations much more consistently.

At moorings near shallow bathymetric features (for example, the Inner Harbour mooring), the modelled semimajor axis is too small and the semi-minor axis is too large, i.e., the current is not focussed enough due to the lack of steering and funnelling by missing features. We anticipate that moving to NEMO4, which has wetting and drying and so can handle these shallow features, will improve the currents in these regions.

The tidal ellipses are not necessarily largest nearest the surface: moorings in the river plume such as Partridge Island have smaller ellipses nearer the surface which increase with depth. SJ100 captures this pattern, which is not visible in CIOPS-E. The other semidiurnal ellipses are small compared to M2, and the diurnal ellipses are even smaller. Figure 36 shows M2, N2, S2 O1, and K1 ellipses for the Partridge Island mooring at 3 m below the surface; other moorings have similar relationships between constituents.

5.2.3.2 Non-tidal circulation

The nontidal circulation is noisy, especially since the processing was done without imposing bandstop filters to remove circulation very near to the M2 frequency. Time series shown are short enough to be able to see low frequency patterns in spite of the high frequency variability. Figure 37 shows the Outer Harbour mooring, which was in the water during the 2019 flood. Both the east/west and north/south velocities increase in magnitude for almost a full month in the spring before and during the flood, and only SJ100 captures this pattern. Consistent with the tidal results, CIOPS-E does not capture the magnitude or variability of the non-tidal observations, and Fundy500 captures only the north-south velocity with reasonable skill. Similar

results are seen for Inner Harbour and Partridge Island moorings that were in the water at the same time, indicating that SJ100 is the most skillful of the three models in the harbour area during the flood. This conclusion is important for e-navigation, as mariners may need to rely more heavily on the dynamic model results when the harbour circulation is in some way unusual.

Time series for near-surface non-tidal velocities for mooring 598 and the Smart Buoy are shown in Figure 38 and Figure 39 respectively. The differences in the signals are more difficult to see, though the histograms show that SJ100 and Fundy500 are still more skillful than CIOPS-E. We use spectral analysis to more clearly isolate the low-to-mid frequency bands; non-tidal rotary spectra plots for 598 and the Smart Buoy are shown in Figure 40 and Figure 41 respectively. CIOPS-E substantively underestimates the signal at almost all frequencies for both deployments. This discrepancy can be over an order of magnitude, especially around the 0.5 cpd frequency. Fundy500 is more skillful than CIOPS-E but tends to underestimate 5–10 cpd frequency band. SJ100 matches the spectral shapes at all frequencies for both of these locations and is the most skillful of the three models. The difference between Fundy500 and SJ100 is not always larger than the 95% confidence interval, but they are typically both an improvement over CIOPS-E at most frequencies.

Figure 42 and Figure 43 show non-tidal rotary spectra for two other deployments: 591 is in the Saint John harbour, and 593 is outside of the SJ100 domain in Mace's Bay. The same patterns are seen in these locations: the port models outperform CIOPS-E (which underestimates the signal at most frequencies), and where SJ100 is available it has more energy at higher frequencies than Fundy500 does. We also note that none of these spectral plots have readily apparent peaks at the near-inertial bands, and we see no substantive difference between the clockwise and counter-clockwise spectra. We conclude that near-inertial oscillations are not prominent in the Saint John harbour area or Mace's Bay, but we have no observations closer to the open boundary regions of Fundy500 to assess whether there are significant levels of near-inertial energy elsewhere in the system.

5.2.3.3 Total velocity

To put the tidal and non-tidal results together, we plot three one-week snippets of total velocity for three different moorings at Canaport, the Smart Buoy, and Partridge Island in Figure 44, Figure 45, and Figure 46 respectively, with a speed / direction decomposition. Only SJ100 is shown for clarity, as previous plots have demonstrated that it outperforms both Fundy500 and CIOPS-E. All three figures show that SJ100 captures the timing of the (largely tidal) signal in both speed and direction, and either matches or slightly underestimates the maximum speed. The direction of the modelled flow is a bit smoother than the observed, and the speed lacks some of the high frequency variability of the observations, but overall, these arbitrarily chosen segments show that SJ100 does a good job of modelling the total velocity in the area of interest.

5.2.3.4 Statistics

Depth profiles of the mean currents and vector correlation of selected ADCPs are shown in Figure 47 to Figure 50. Consistent with the rest of the analysis, SJ100 outperforms Fundy500 and CIOPS-E in matching the observations, especially near the surface. Both total and non-tidal mean currents are shown, and SJ100 consistently matches the shape of the mean speed profile, even if it does not quite match the amplitude of the profile. The magnitude of the vector correlation varies significantly with depth, and we see higher correlations higher in the water column. The correlation angle is also plotted, although any interpretation of this is robust only when the amplitude of the correlation is high.

Moorings are often deployed in groups, and so we can aggregate depth profiles of statistics for 3 or 4 concurrent moorings. Harbour moorings from 2016 are shown in Figure 51, harbour moorings in 2019 are shown in Figure 52, and a spread of moorings from Mace's Bay to Cape Spencer is shown in Figure 53. In each of these plots, the top row shows the mean bias for u and v velocities, and the middle and bottom rows show the CRMSE and γ^2 of total, tidal, and non-tidal currents. Not all depths considered had enough data to do a tidal analysis, and so some of the dashed and dotted lines in the middle and bottom row are more limited than the solid lines. Across the three plots we see that the port model outperforms CIOPS-E in the Saint John harbour, but the improvement is not as large farther away from the harbour. Bias is most strongly improved with increased resolution near the harbour. Total CRMSE is dominated by the tidal component, which suggests that to further improve the results, we should focus on the tidal circulation patterns. In contrast, γ^2 values are smaller for tidal and total circulation than they are for non-tidal circulation. SJ100 has the smallest (i.e., best) values of γ^2, although $\gamma^2 > 1$ for non-tidal circulation at some moorings.

5.2.4 ADCP transects

Seven ADCP transects were conducted around the Saint John harbour in the spring of 2019, and three representative transects are shown in Figure 54, Figure 55, and Figure 56 comparing the observations with model results. One transect crosses the river where it meets the harbour, one transect starts at the river-harbour junction and extends just beyond the Partridge Island Sea wall, and another crosses the river plume just south of Partridge Island. In all plots, v is the cross-transect velocity and u is the along-transect velocity, and all three transects were conducted on the same day (April 26, 2019). SJ100 captures most of the circulation patterns: the amplitudes are typically comparable or underestimated, particularly when around some of the shallow features of the harbour which steer and focus circulation patterns. The transect that crosses the mouth of the river roughly connects the two HADCP locations, and so is only present in the SJ100 domain. Importantly, it shows cross-channel variability that we do not capture in our (still relatively coarse) model, which is consistent with results from the HADCPs. There is marine traffic that crosses this transect: if the cross-channel variability is important to mariners, this may add motivation to increasing the resolution in the harbour.

To compare transects from all three models, Figure 57 shows a scatter plot of bias vs CRMSE. Each transect has a different marker shape, and the model results are indicated by colour. Better results are found closer to the origin, and SJ100 is consistently more skillful at matching the currents than either Fundy500 or CIOPS-E. The scores are taken as an aggregate of all points at all depths, not just the near-surface velocities, and so provide a measure for how well the vertical structure of the horizontal velocity is captured.

5.3 WATER PROPERTIES

Long term surface temperature time series are available across the Bay of Fundy, at buoys, NOAA tide gauge stations, and a research station in upper Fundy (see Figure 58). Temperature and salinity measurements from moored CTDs are available largely in the Saint John River and surrounding intertidal areas (see Figure 64). CTD casts are available in Canadian waters throughout the domain, though most casts are taken near Saint John or the Passamaquoddy area early in the hindcast period.

5.3.1 Sea surface temperature

There are three buoys with sea surface temperature available: the Smart Atlantic buoy (Figure 59), the ECCC-deployed buoy West Bay of Fundy (44490, Figure 60), and a buoy run by Fundy FORCE in Minas Basin (Figure 61). There are also temperature time series available at the two

NOAA gauges in the domain (Figure 62 and Figure 63). The Smart Atlantic buoy is the only buoy in the SJ100 domain and is the only buoy that spans the entire hindcast period: the West Fundy buoy was installed in 2019 (and lost in June 2022), and the FORCE buoy was also installed in 2019.

All three models (CIOPS-E, Fundy500, and SJ100) capture the seasonal cycle of SST, matching both the timing of the seasonal shifts and the seasonal highs and lows. There is a warm bias of a degree or two in the summer months at the West Fundy buoy in particular; this warm bias is partially, but not entirely, inherited from CIOPS-E, and may be improved by adjusting the light penetration parameters. At the Smart Buoy, Fundy500 and SJ100 do not fully capture the range of variability of the SST. This buoy is close enough to the harbour that the SST is influenced by the temperature of the river water, and since we use a climatological river temperature, we are missing the high-frequency variability signal from the river plume. At the FORCE gauge, both CIOPS-E and Fundy500 capture the SST signal quite well, with the exception of occasional spikes. It is unclear whether these spikes are due to cold water intrusions that are not modelled or are artifacts of the buoy being sloshed around and tipping into the air. While the upper Fundy region is not a focus of this model development, it is encouraging that our Fundy500 SST signal matches observations well.

These qualitative comparisons are supported by the statistics in Table 22: at the Smart Atlantic buoy, SJ100 outperforms CIOPS-E and Fundy500 (except for the overall bias, where Fundy500 has a smaller value). The bias at all stations is under 2°C, and less than half a degree at the Smart Buoy. The two stations at the coast show larger biases than buoys moored offshore, suggesting that this may be due to the representation of the atmosphere at the land-sea interface. CRMSE is around 1–1.5° for most stations, and time series show that it is mostly high-frequency variability that we are missing. The γ^2 values are under 0.1, indicating a good level of skill in the port models, and Fundy500 and SJ100 outperform CIOPS-E.

5.3.2 Moored CTDs

Moored CTDs (MCTDs) were deployed mostly near the coast as part of the Coastal Environmental Baseline Program, and only instruments that are in open water are considered in this evaluation (see Figure 64). Since our system does not include freshwater runoff from rivers other than the Saint John River, we do not fully capture the observed salinity variation in smaller estuaries along the coast (in particular, in the Musquash estuary MPA). Statistics for all MCTDs are shown in Table 23. CIOPS-E does not include the Saint John River, and the Musquash moorings are so close to shore that they are on land in the CIOPS-E domain.

Three deployments were done at the headlands of the Musquash estuary, where the instruments are located near the sea bed at locations exposed to the open bay (see Figure 65, Figure 66, and Figure 67). We generally capture the low frequency signal quite well, though we miss a persistent pulse of fresher water in the fall due to lacking any freshwater runoff from the (small) Musquash River. The fall freshet is both weaker and longer in duration than the typical fall freshets seen in the Saint John River signals, possibly because the river outflow is not strong enough to push the freshwater signal much beyond the extent of the estuary. The observations have more high frequency variability than the port model configurations do, though this again is likely due to the lack of gauged runoff. Minimal contemporary gauge data is available for the river, and since the region includes intertidal salt marshes it may be difficult to acquire. It is encouraging that even without any Musquash river outflow we capture as much of the temperature and salinity signals at the estuary headlands as we do.

There were also instruments deployed in the Saint John river (labelled “The Narrows”) and Courtenay Bay, and two near-surface deployments are shown in Figure 68 and Figure 69. The

Narrows deployment was just upstream of Reversing Falls, and both SJ100 and Fundy500 capture the low frequency variability and the timing of the tidal oscillations well. The high frequency oscillations between salty and fresh water are not fully captured, and the models may be slightly overmixing at Reversing Falls, or just not capturing the sharpness of the edge between fresh and salty water. A Courtenay Bay deployment is included even though it was located in an intertidal area because it is an important area for e-navigation due to the oil terminals located at the head of the bay. With the exception of the cold spikes, we capture the temperature signal well, though we do not capture the variability or magnitude of the salinity.

There is also a deep-water temperature time series near Grand Manan, outside of the SJ100 domain (see Figure 70). The Fundy500 signal matches CIOPS-E quite closely, but like CIOPS-E is too warm in the summer and slightly too cold in the fall. There are very few observations available at depth outside of the area of interest, so it is unclear whether this is indicative of a persistent bias and whether or not this bias is inherited from the CIOPS-E boundary conditions.

5.3.3 CTD profiles

The Fundy500 domain is divided into five regions for CTD cast analysis, shown in Figure 71: the Saint John region (SJ), Passamaquoddy Bay (Pmqdy), and Upper, Middle, and Lower Fundy ({U,M,L}Fundy). While the Saint John and Passamaquoddy regions are relatively easy to determine, the divisions between Upper, Middle, and Lower Fundy are somewhat arbitrary. The Saint John region includes some casts that were taken in the river system, though none are upstream of Reversing Falls. Additionally, there are no observations used that were gathered in US waters, as all sources of CTD casts used were Canadian; this only impacts the Lower Fundy region. Figure 71 plots cast locations used over the entire hindcast with in red dots. Depth profiles of temperature and salinity bias and CRMSE for plots for the Saint John region for 2016 and 2019 are shown in Figure 72 and Figure 73, and for all five regions over the entire hindcast from Figure 74 through Figure 78.

In Saint John, the port models have a slightly larger temperature bias than CIOPS-E, and the port models are warmer than observations by about 0.5°C at all depths over the course of the entire hindcast. CIOPS-E has nearly no bias at depth and is slightly too cold near the surface. The CRMSE values are relatively large, and the bias ± CRMSE intervals heavily overlap for all three models. The port models outperform CIOPS-E for salinity, with smaller biases and smaller CRMSEs. CIOPS-E uses a climatological runoff, and so cannot capture the variability of the near-surface salinity signal. These patterns are broadly similar in the Middle Fundy region as well, though CIOPS has a larger temperature bias than it did in the Saint John region. SJ100 slightly outperforms Fundy500 on temperature, and all three models are comparable on salinity.

Most of the casts used in this evaluation were gathered in 2016, so the summary plot for the Saint John regions looks quite similar for year2016 as it does for the entire hindcast period. The 2019 year had the next the greatest number of casts available, and the temperature patterns are similar to those in 2016 and the hindcast as a whole. The salinity signal in 2019 has a larger large salty bias (~ 8 psu at most) than in 2016 (nearly no net bias for SJ100). There are about 50 casts in the region in 2016, but only about 20 in 2019, and some of the 2019 fieldwork was done during the spring freshet and flood. We likely underestimate the freshwater content of the near-surface harbour water during the flood.

The other three regions are covered only by Fundy500 and CIOPS-E. In both the Lower Fundy and Passamaquoddy regions, Fundy500 shows a +0.5–1°C warm bias, while CIOPS-E has about bias of only ~ +0.25–0.5° C. The salinity signals are comparable between Fundy500 and CIOPS-E in Lower Fundy, with a salty bias of about +1 psu and a large CRMSE. The signals differ more significantly in Passamaquoddy, though the large CRMSE values for temperature in

CIOPS-E suggest that resolution may be an issue for CIOPS-E in the region. There are not a large number of casts in the Upper Fundy region, but Fundy500 matches CIOPS-E in temperature and has a smaller CRMSE for salinity. There are sources of fresh water in the upper Fundy region that are not included in Fundy500 but are accounted for in the runoff climatology used in CIOPS-E.

5.3.4 Ferries

There are no ferries fitted with thermosalinographs in the Bay of Fundy, though there are longstanding discussions to outfit the Digby-to-Saint John ferry with one.

5.4 DRIFT

Drifter data used for evaluation in the Bay of Fundy and Saint John models comes from two sources. Fourteen Davis drifters were released in Saint John Harbour between December 2019 and March 2020 as part of the OPP program, in support of model evaluation (Schillinger). They were released in pairs approximately biweekly over this period. A second set of 108 drifters were deployed by the Fundy North Fisherman's Association (FNFA) as part of OPP's Coastal Environmental Baseline Program (Mitchell, 2022). These drifters were designed at the Saint Andrews Biological Station (SABS) and take two forms. Half the FNFA drifters are Disk drifters, which are circular, 50 cm in diameter and 11 cm height, floating halfway submerged. The other half are Long drifters, 72.5 cm high and 11.5 cm wide, comprised of sets of vertically oriented orthogonal fins. The FNFA drifters were released in the harbor near the Canaport LNG terminal in tight clusters of 4–5 drifters of the same type, approximately monthly throughout 2021. The paths of all drifters used for model evaluation are shown in Figure 79.

The drag area ratio (DAR) and windage for the SABS-designed drifters were calculated following the same method as applied to commercially available drifter types in Hourston (2021). Drag area ratios of 0.034 and 0.236 for the SABS Long and Disk drifters yield effective windages of 1.64 and 2.68%, which include 1% for Stokes drift, as recommended by Sutherland et al (2020). The Davis drifter type has a DAR of 0.011 (Hourston, 2021) and windage of 1.27%.

All Saint John area drifter data were interpolated onto an hourly time base and cropped to be within the land and open ocean boundaries of the SJ100 model domain. Drift simulations were initiated each hour at the position of the real drifters and run for 25 hours. Drift simulations were carried out using surface layer currents from the SJ100, Fundy500, and CIOPS-E models, which are all approximately 1 m thick.

Figure 80, Figure 81 and Figure 82 show the mean Molcard skill score (top) and mean separation distance (bottom) of Davis, Long, and Disk drifters, respectively, as a function of simulation hour for each circulation model. Scores at 4-hourly time points are reported in Table 24. For the Davis drifters, all three models perform similarly well, while for the Long and Disk drifters, the CIOPSE score is lower than those of Fundy500 and SJ100 models. For all three drifter types, the three models' results all fall well within the range of one standard deviation. The mean separation distances increase continuously over the 25-hour simulation periods, with a faster increase earlier in the simulations, and a more gradual increase later in the periods. The mean Molcard skill scores decline over the simulation period, but also show a strong periodic signal at the dominant M2 tidal period. The periodic behavior of the Molcard score is explained by its dependence on the distance from the drifters' initial position in the denominator of the score calculation (see section 4.1.4). In this strongly tidal area, the dominant motion of each drifter follows a tidal orbital. Regardless of starting time, each drifter returns to near its starting position after completing a full tidal ellipse. At these times, the drifter's displacement from the starting position is small compared to its displacement at times just before or after a complete

tidal period. Separation distances increase monotonically, so the Molcard score's variation with M2 period indicates the Molcard score may not be the best choice for a strongly tidal flow regime.

The Disk drifters (Figure 82) simulation skill is lower than that of the Davis (Figure 80) and Long (Figure 81) drifters for the first 12 hours and similar during the second 12 hours. A lower skill is expected because the Disks only occupy the top 5.5 cm of the water column, compared to 0.95 m and 0.725 m for Davis and Long, which are much closer to the top ~1 m thickness represented by the model surface layer. This result indicates that either the surface 1 m currents are vertically uniform, or that any shear in the top 1 m is proportional to the wind and has been reasonably accommodated by the larger windage used for Disk drifters.

Many factors may contribute to the large standard deviations of skill scores and separation distances, which indicate great variation in space and time of the accuracy of the drift prediction. Besides the potential inaccuracies of the three circulation models, the effect of wind and waves on these drifters is poorly constrained. For the SABS drifters in particular, the windage is estimated from the shape of the drifters, but no field studies have been published to establish their typical windages. For all three types of drifters, the extra windage added to account for Stokes drift is taken from literature based on open-ocean analyses, and not based on any information specific to the wave field near Saint John, which is very sheltered from open ocean swell and subject to a short fetch from northerly and westerly winds. As such, there may be substantial room to improve skill at drift simulation if simulation parameters were optimized for models with high horizontal resolution and tuned for this geographic region.

6 FORECAST EVALUATION RESULTS

Forecast evaluation is done over a two month period from December 3, 2021, to January 31, 2022, and 48-hour-long forecasts are launched daily from the daily pseudo-analysis (i.e., hindcast). Only water level and SST observations are used to evaluate the forecasts, as there were no usable HADCP observations available in the two month window. Non-tidal water level is evaluated at Saint John (00065), Eastport (NOAA gauge 8410140) and Cutler Farris (NOAA gauge 8411060). SST is evaluated at the Smart Buoy, buoy 44490, FORCE, Eastport, and Cutler Farris.

6.1 NON-TIDAL WATER LEVEL

The bias and CRMSE of non-tidal water level are plotted as functions of lead hour for each of the three tide gauges in Figure 83, Figure 84, and Figure 85. The bias for Fundy500 at Cutler Farris and Eastport are about -0.03 m and -0.17 m respectively, and CIOPS-E has produced slightly more skillful results; there is a semidiurnal signal visible at Cutler Farris, but no obvious trend at Eastport over the 48 hours. Both port model domains outperform CIOPS-E at Saint John, and SJ100 has a smaller bias (-0.02 m) and smaller CRMSE (~0.07 m) than Fundy500. The variability of CRMSE is slightly larger towards the end of the forecast than at the beginning, though the ensemble average stays smaller than 0.1m for SJ100 at all lead hours.

6.2 HORIZONTAL ADCP

No useable HADCP data was available during the forecast evaluation period.

6.3 SEA SURFACE TEMPERATURE

Plots of bias and CRMSE as a function of lead hour for at buoy 44490, the Smart Buoy, and FORCE are plotted in Figure 86, Figure 87, and Figure 88. These three locations span most of the domain, giving a good overview of how SST is modelled through the region.

Buoy 44490 is located off Grand Manan, near the entrance to the Bay of Fundy. As with the hindcast, CIOPS-E has a smaller (warm) bias than Fundy500 at all lead hours, though Fundy500 is no more than half a degree warmer than observations. However, Fundy500 improves on CIOPS-E's CRMSE, perhaps in part because the tidal circulation is improved in Fundy500 relative to CIOPS-E. At the Smart Buoy near Saint John, CIOPS-E is slightly warm, Fundy500 is slightly cold, and SJ100 has no substantial net bias. The CRMSE is also smaller in SJ100 than either Fundy500 or CIOPS-E, indicating that the variability is better captured. Given that the river plume lacks a directly forced high-frequency temperature signal from the upstream river water, these results are encouraging. The FORCE results show a clear semidiurnal signal, indicating that there are discrepancies in the tidal circulation in the area; this is not unexpected.

7 SUMMARY

We have shown that the Fundy500 and SJ100 port models are robust, stable, and skillful. We ran a six year hindcast that included two historic floods and four hurricanes, and the model remained stable and produced physical results throughout. We also ran two months of 48 hour long forecasts in the winter, with forecasts launched daily.

In the immediate Saint John area, we find that SJ100 is the most precise at reproducing observations of the three model solutions considered; this is particularly true for water level and currents. Long term observations including the tide gauges and the Smart Atlantic buoy show that the models perform consistently from year to year. Observations are concentrated in the Saint John area, but both Fundy500 and CIOPS-E can reproduce observations in the lower and middle parts of the Bay of Fundy as well. Tidal results in the upper Fundy area are unexpectedly decent for Fundy500, and we anticipate that they will be further improved by moving to NEMO4. SJ100 is consistently the best at reproducing observed currents in the Saint John area, at times far outperforming CIOPS-E.

Our drift analysis shows that the port model performs comparably to CIOPS-E in a statistical sense. Fundy500 and SJ100 do have slightly higher Molcard scores than CIOPS-E for the Long drifters, but slightly lower scores for Davis drifters and nearly the same scores for the Disk drifters. The standard deviation of both statistics is substantial for all three drifter types, and the differences between the models is not significant.

The two months of forecasts were also analyzed. We find no significant growth in error over the course of the 48 hour forecast length, and the port models are skillful at reproducing total and non-tidal water level as well as SST. The port models give more skillful results than CIOPS-E, with smaller biases and CRMSE.

We conclude that the two-level port model system (Fundy500 and SJ100) is a net improvement over the existing model solution (CIOPS-E). Our evaluation was designed with the applications of drift prediction and e-navigation in mind, and we have shown that in particular, the port model is skillful at reproducing near-surface currents, water level, and surface temperature.

8 KEY FINDINGS

1. SJ100 and Fundy500 outperform or match the skill of CIOPS-E for all types of observations considered.

2. SJ100 gives more accurate tidal and non-tidal water level results than either Fundy500 or CIOPS-E at Saint John. Fundy500 provides improved results over CIOPS-E.
3. SJ100 is robustly the most skillful of the three domains considered at reproducing the currents in and around the Saint John harbour. It accurately reproduces both tidal and non-tidal circulation better than either Fundy500 or CIOPS-E, especially in the inner harbour.
4. HADCP results show that the river outflow is adequately represented by SJ100, but is not properly captured by either Fundy500 or CIOPS-E.
5. SST and MCTD variability is well captured by SJ100, though all three models show a small warm bias.
6. SJ100 and Fundy500 have a warm bias in CTD casts that is not seen in CIOPS-E, though the CRMSE values are nearly as large as the bias. Both SJ100 and Fundy500 show better salinity results than CIOPS-E.
7. Drift results for Fundy500 and SJ100 are comparable to those for CIOPS-E.
8. Forecasts of water level and surface temperature do not show substantive growth in bias or CRMSE over the course of the forecast. SJ100 provides best results for surge, total water level, and surface temperature in the Saint John area.

9 ACKNOWLEDGEMENTS

The original Saint John port model on which this POPS is based was developed by Jean-Philippe Paquin, Youyu Lu, and Xianmin Hu, with additional support from Mitchell O'Flaherty Sproul and Susan Haigh. Observations were provided and processed by Doug Schillinger, Fred Page, Peter Kraska, Justine McMillan, the Coastal Environmental Baseline Program, the Fundy North Fisherman's Association, and Fundy FORCE. Bathymetry and coastline data provided by the Canadian Hydrographic Service and Phil MacAulay. The Matlab code for NS_TIDE was provided by Pascal Matte. The drift analysis was done using tools built by Nancy Soontiens and Jennifer Holden. All runs were conducted using a maestro suite, and maestro support was provided by Sarah MacDermid and Ji Lei. Many thanks to all these colleagues and collaborators for their contributions and insight.

10 REFERENCES CITED

Aretxabaleta, Alfredo L., Dennis J. McGillicuddy, Keston W. Smith, and Daniel R. Lynch. 2008. "Model Simulations of the Bay of Fundy Gyre: 1. Climatological Results." Journal of Geophysical Research 113 (C10): C10027. https://doi.org/10.1029/2007JC004480.

Aretxabaleta, Alfredo L., Dennis J. McGillicuddy, Keston W. Smith, James P. Manning, and Daniel R. Lynch. 2009. "Model Simulations of the Bay of Fundy Gyre: 2. Hindcasts for 2005–2007 Reveal Interannual Variability in Retentiveness." Journal of Geophysical Research 114 (C9): C09005. https://doi.org/10.1029/2008JC004948.

Backhaus, Jan O. 1983. "A Semi-Implicit Scheme for the Shallow Water Equations for Application to Shelf Sea Modelling." Continental Shelf Research 2 (4): 243–54. https://doi.org/10.1016/0278-4343(82)90020-6.

———. 1985. "A Three-Dimensional Model for the Simulation of Shelf Sea Dynamics." Deutsche Hydrographische Zeitschrift 38 (4): 165–87. https://doi.org/10.1007/BF02328975.

Blanken, Hauke, Caterina Valeo, Charles Hannah, Usman T. Khan, and Tamás Juhász. 2021. “A Fuzzy-Based Framework for Assessing Uncertainty in Drift Prediction Using Observed Currents and Winds.” Frontiers in Marine Science 8 (November): 618094. https://doi.org/10.3389/fmars.2021.618094.

Chen, Changsheng, Hedong Liu, and Robert C. Beardsley. 2003. “An Unstructured Grid, Finite-Volume, Three-Dimensional, Primitive Equations Ocean Model: Application to Coastal Ocean and Estuaries.” Journal of Atmospheric and Oceanic Technology 20 (1): 159–86. https://doi.org/10.1175/1520-0426(2003)020<0159:AUGFVT>2.0.CO;2.

Cummins, Patrick F., and Lie-Yauw Oey. 1997. “Simulation of Barotropic and Baroclinic Tides off Northern British Columbia.” Journal of Physical Oceanography 27 (5): 762–81. https://doi.org/10.1175/1520-0485(1997)027<0762:SOBABT>2.0.CO;2.

Cummins, Patrick F., and Pramod Thupaki. 2018. “A Note on Evaluating Model Tidal Currents against Observations.” Continental Shelf Research 152 (January): 35–37. https://doi.org/10.1016/j.csr.2017.10.007.

Daniel, Pierre, Gwénaële Jan, Fanch Cabioc’h, Yves Landau, and Erwann Loiseau. 2002. “Drift Modeling of Cargo Containers.” Spill Science & Technology Bulletin 7 (5–6): 279–88. https://doi.org/10.1016/S1353-2561(02)00075-0.

Dupont, Frédéric, Charles G. Hannah, and David Greenberg. 2005. “Modelling the Sea Level of the Upper Bay of Fundy.” Atmosphere-Ocean 43 (1): 33–47. https://doi.org/10.3137/ao.430103.

Dupont, Frédéric, Simon Higginson, Romain Bourdallé-Badie, Youyu Lu, François Roy, Gregory C. Smith, Jean-François Lemieux, Gilles Garric, and Fraser Davidson. 2015. “A High-Resolution Ocean and Sea-Ice Modelling System for the Arctic and North Atlantic Oceans.” Geoscientific Model Development 8 (5): 1577–94. https://doi.org/10.5194/gmd-8-1577-2015.

Garrett, Christopher. 1972. “Tidal Resonance in the Bay of Fundy and Gulf of Maine.” Nature 238 (5365): 441–43. https://doi.org/10.1038/238441a0.

Horwitz, Rachel M., Stephanne Taylor, Youyu Lu, Jean-Philippe Paquin, Doug Schillinger, and David A. Greenberg. 2021. "Rapid reductino of tidal amplitude due to form drag in a narrow channel." Continental Shelf Research, 213, 104299. doi:10.1016/j.csr.2020.104299

Hourston, Roy A.S. 2021. Surface Ocean Circulation Tracking Drifter Data from the Northeastern Pacific and Western Arctic Oceans, 2014-2020. Sidney, B.C.: Fisheries and Oceans Canada - Pêches et Océans Canada.

Ku, L.-F., D. A. Greenberg, C. J. Garrett, and F. W. Dobson. 1985. "Nodal Modulation of the Lunar Semidiurnal Tide in the Bay of Fundy and Gulf of Maine." Science, 230, 69-71. https://www.science.org/doi/10.1126/science.230.4721.69.

Kundu, Pijush K. 1976. “Ekman Veering Observed near the Ocean Bottom.” Journal of Physical Oceanography 6 (2): 238–42. https://doi.org/10.1175/1520-0485(1976)006<0238:EVONTO>2.0.CO;2.

Large, William George, and Stephen G. Yeager. 2004. “Diurnal to Decadal Global Forcing for Ocean and Sea-Ice Models: The Data Sets and Flux Climatologies.” NCAR Technical Note NCAR/TN-460+STR. Boulder, CO: Climate and Global Dynamics Division: National Center for Atmospheric Research.

Madec, Gurvan. 2016. NEMO Ocean Engine. Note Du Pole de Modélisation de l’Institut Pierre-Simon Laplace 27. NEMO European Consortium.

Matte, Pascal., Jay, David A., and Edward D. Zaron. 2013. Adaptation of Classical Tidal Harmonic Analysis to Nonstationary Tides, with Application to River Tides. Journal of Atmospheric and Oceanic Technology, 30(3), 569-589.

Milbrandt, Jason A., Stéphane Bélair, Manon Faucher, Marcel Vallée, Marco L. Carrera, and Anna Glazer. 2016. “The Pan-Canadian High Resolution (2.5 Km) Deterministic Prediction System.” Weather and Forecasting 31 (6): 1791–1816. https://doi.org/10.1175/WAF-D-16-0035.1.

Mitchell, L. 2022. Data provided by the Fundy North FIsherman's Association. Personal Communication. Retrieved from www.fundynorth.org/opp

Molcard, A., P.M. Poulain, P. Forget, A. Griffa, Y. Barbin, J. Gaggelli, J.C. De Maistre, and M. Rixen. 2009. “Comparison between VHF Radar Observations and Data from Drifter Clusters in the Gulf of La Spezia (Mediterranean Sea).” Journal of Marine Systems 78 (November): S79–89. https://doi.org/10.1016/j.jmarsys.2009.01.012.

Niiler, Pearn P., Andrew S. Sybrandy, Kenong Bi, Pierre M. Poulain, and David Bitterman. 1995. “Measurements of the Water-Following Capability of Holey-Sock and TRISTAR Drifters.” Deep Sea Research Part I: Oceanographic Research Papers 42 (11–12): 1951–64. https://doi.org/10.1016/0967-0637(95)00076-3.

Nudds, Shannon, Youyu Lu, Simon Higginson, Susan P. Haigh, Jean-Philippe Paquin, Mitchell O’Flaherty-Sproul, Stephanne Taylor, et al. 2020. “Evaluation of Structured and Unstructured Models for Application in Operational Ocean Forecasting in Nearshore Waters.” Journal of Marine Science and Engineering 8 (7): 484. https://doi.org/10.3390/jmse8070484.

Paquin, Jean-Philippe, Youyu Lu, Stephanne Taylor, Hauke Blanken, Guillaume Marcotte, Xianmin Hu, Li Zhai, et al. 2020. “High-Resolution Modelling of a Coastal Harbour in the Presence of Strong Tides and Significant River Runoff.” Ocean Dynamics 70 (3): 365–85. https://doi.org/10.1007/s10236-019-01334-7.

Paquin, Jean-Philippe, François Roy, Gregory C. Smith, Frédéric Dupont, Sarah MacDermid, Yukie Hata, Oleksandr Huizy, Ji Lei, Yosvany Martinez, Hauke Blanken, Jennifer Holden, Nancy Soontiens. 2021a. "Coastal Ice Ocean Prediction System for the East Coast of Canada (CIOPS-E): Update from version 1.50 to 2.0.0." Canadian Centre for Meteorological and Environmental Prediction Tech. Note. https://collaboration.cmc.ec.gc.ca/cmc/cmoi/product_guide/docs/tech_notes/technote_ciops-east-200_e.pdf

Paquin, Jean-Phillipe, Gregory C. Smith, Frederic Dupont, Sarah MacDermid, Yukie Hata, Oleksandr Huizy, Ji Lei, Yosvany Martinez, Michael Dunphy, Maxim Krassovski, Stephanne Taylor, Hauke Blanken, Jennifer Holden, Nancy Soontiens, Susan E. Allen, and Doug Latornell. 2021b. “Coastal Ice Ocean Prediction System for the West Coast of Canada (CIOPS-W): Update from version 1.5.0 to 2.0.0.” Canadian Centre for Meteorological and Environmental Prediction Tech. Note. https://collaboration.cmc.ec.gc.ca/cmc/cmoi/product_guide/docs/tech_notes/technote_ciops-west-200_e.pdf.

Paquin, Jean-Phillipe, Gregory C. Smith, Francois Roy, Frederic Dupont, Sarah MacDermid, Ji Lei, Youyu Lu, Stephanne Taylor, Hauke Blanken, and Li Zhai. 2022. “Coastal Ice Ocean Prediction System for the West Coast of Canada (CIOPS-W): System Description for Versions 1.0.0 and 1.5.0.” Canadian Centre for Meteorological and Environmental Prediction Tech. Note. https://collaboration.cmc.ec.gc.ca/cmc/cmoi/product_guide/docs/tech_notes/technote_ciops-west-100_e.pdf.

Robin, Catherine, Shannon Nudds, Phillip M. MacAulay, André G. Godin, Bodo De Lange Boom, and Jason Bartlett. 2016. "Hydrographic Vertical Separation Surfaces (HyVSEPs) for the Tidal Waters of Canada." Marine Geodesy, 39(2), 195-222. doi:10.1080/01490419.2016.1160011

Röhrs, Johannes, and Kai H. Christensen. 2015. “Drift in the Uppermost Part of the Ocean.” Geophysical Research Letters 42 (23). https://doi.org/10.1002/2015GL066733.

Röhrs, Johannes, Kai Håkon Christensen, Lars Robert Hole, Göran Broström, Magnus Drivdal, and Svein Sundby. 2012. “Observation-Based Evaluation of Surface Wave Effects on Currents and Trajectory Forecasts.” Ocean Dynamics 62 (10–12): 1519–33. https://doi.org/10.1007/s10236-012-0576-y.

Saucier, François J. 2003. “Modeling the Formation and Circulation Processes of Water Masses and Sea Ice in the Gulf of St. Lawrence, Canada.” Journal of Geophysical Research 108 (C8): 3269. https://doi.org/10.1029/2000JC000686.

Saucier, François J., and Joël Chassé. 2000. “Tidal Circulation and Buoyancy Effects in the St. Lawrence Estuary.” Atmosphere-Ocean 38 (4): 505–56. https://doi.org/10.1080/07055900.2000.9649658.

Schillinger, Doug. J. (n.d.). Oceanographic data collected at Saint John Harbour in support of the development of a real-time, port scale, prediction system. in preparation.

Smith, Stuart D. 1988. “Coefficients for Sea Surface Wind Stress, Heat Flux, and Wind Profiles as a Function of Wind Speed and Temperature.” Journal of Geophysical Research 93 (C12): 15467. https://doi.org/10.1029/JC093iC12p15467.

Soontiens, Nancy, and Jennifer Holden J. 2024. "A framework for evaluation and characterization of drift prediction in the marine environment." Can. Tech. Rep. Hydrogr. Ocean Sci. 383: ix + 45 p. https://waves-vagues.dfo-mpo.gc.ca/library-bibliotheque/41252287.pdf.

Sutherland, Graig, Nancy Soontiens, Fraser Davidson, Gregory C. Smith, Natacha Bernier, Hauke Blanken, Douglas Schillinger, et al. 2020. “Evaluating the Leeway Coefficient of Ocean Drifters Using Operational Marine Environmental Prediction Systems.” Journal of Atmospheric and Oceanic Technology 37 (11): 1943–54. https://doi.org/10.1175/JTECH-D-20-0013.1.

Zhang, Minghong, William Perrie, and Zhenxia Long. 2019. “Springtime North Pacific Oscillation and Summer Sea Ice in the Beaufort Sea.” Climate Dynamics 53 (1–2): 671–86. https://doi.org/10.1007/s00382-019-04627-1.

11 TABLES

Table 1: Key model setup parameters Fundy500 and SJ100.

Namelist parameter	Outer Grid (Fundy500)	Inner Grid (SJ100)
Number of grid points	929 x 429 x 40	799 x 1009 x 34
Horizontal Resolution	~420 m	~85 m
Vertical Resolution	1 m – 18 m	1 m – 16 m
Time step (baroclinic / barotropic)	30 s / 1 s	10 s / 0.333 s
OBC SSH offset	0.69 m (only S and W)	0.0 cm
Source of tidal forcing	WebTide (Scotia-Fundy-Maine) and inference	Fundy500
Tidal constituents	M2, S2, N2, K1, O1, L2, K2, NU2, 2N2, M4, P1, Q1, λ2, T2	M2, S2, N2, K1, O1, L2, K2, NU2, 2N2, M4, P1, Q1, λ2, T2
Equation of State	EOS-80	EOS-80
Free Surface	Variable volume	Variable volume
Light penetration	Two band	Two band
Lateral boundary condition	Partial slip: shlat = 1	Partial slip: shlat = 1
Momentum Advection	Vector form, with 5 substeps for vertical advection	Vector form, with 5 substeps for vertical advection
Momentum Lateral Diffusion	Horizontal Laplacian and Smagorinsky	Horizontal Laplacian and Smagorinsky
Tracer Advection	Total Variance Dissipation, with 5 vertical substeps	Total Variance Dissipation, with 5 vertical substeps
Tracer Lateral Diffusion	Iso-neutral Laplacian and Smagorinsky	Iso-neutral Laplacian and Smagorinsky
Vertical diffusion	k-ε (GLS)	k-ε (GLS)
Bottom friction	Quadratic, with additional friction at Rev. Falls and Minas Passage.	Quadratic, with additional friction at Rev. Falls

Table 2: Inferred tidal constituents used in Fundy500. First constituent of each pair is inferred, and it is inferred relative to the second constituent.

Constituents	Amplitude Ratios (Yarmouth, Cutler-Farris)	Phase Differences (Yarmouth, Cutler-Farris)	Amplitude	Phase
λ2 and L2	0.301, 0.370	-5, -23	0.335 m	-14.0°
P1 and K1	0.300, 0.334	0, +2	0.315 m	+1.0°
Q1 and O1	0.157, 0.159	-14, -19	0.158 m	-16.5°
T2 and S2	0.089, 0.087	-20, -21	0.088 m	-20.5°

Table 3: Tuning adjustments to tidal constituents between Fundy500 and SJ100.

Constituent	Amplitude	Phase	Constituent	Amplitude	Phase
S2	-0.04 m	-4.0°	K2	None	-6.0°
N2	None	-4.0°	λ2	None	+7.5°
O1	None	+5.5°	P1	None	-132.0°
K1	None	+5.5°	T2	+0.01 m	-23.5°
L2	+ 0.01 m	+18.0°	Q1	None	-62.5°
2N2	- 0.02 m	-29.0°	M4	None	-252.0°
NU2	+0.01 m	+6.0°	—	—	—

Table 4: Description of gauges used in NS_TIDE calculations.

Gauge ID	Type of signal	Used As	Lag in NS_TIDE	Fallback procedure
01AP003 – Oak Point	Mixed fluvial–tidal	Hindcast OBC, Forecast OBC	—	HC: Substantial gaps patched with NS_TIDE Forecast: Predicted using NS_TIDE
01AK003 – Fredericton	Fluvial	NS_TIDE upstream gauge	+ 18.5 hours	Persistence, then a gradual fallback to climatology
00065 – Saint John	Tidal	NS_TIDE downstream gauge	0 hours	None needed, as data used is static constituents calculated from observations
01AP005 – Reversing Falls	Highly mixed fluvial–tidal	NS_TIDE gauge to include phase lag over sill	+ 2 hours	Harmonic constituents (to avoid requiring a recursive NS_TIDE solution)

Table 5: Vertical datum conversion values for water level gauges.

Station	Original Datum	Conversion to CGVD28
00065 – Saint John	Chart Datum	-4.17 m
8410140 – Eastport	Mean low-low water	-2.69 m
8411060 – Cutler Farris	Mean low-low water	-2.07 m
00320 – Parker's Cove	Chart Datum	-4.88 m

Table 6: M2 water level tidal parameters for 2017. Tidal error and amplitude units are metres, and phase units are degrees.

Observation			Amplitude Error			Phase Error			Tidal Error		
Station name	**Ampl**	**Phase**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**
Cutler Farris	2.022	93.61	+0.128	+0.035	-	-1.33	+1.60	-	0.097	0.047	-
Eastport	2.654	98.27	+0.336	+0.007	-	+2.78	+3.23	-	0.257	0.106	-
Saint John	3.055	97.92	+0.249	+0.023	+0.037	+0.01	-0.72	-0.44	0.176	0.032	0.031
Parkers Cove	3.433	88.69	+0.221	-0.141	-	+3.10	+1.48	-	0.207	0.117	-

Table 7: N2 water level tidal parameters for 2017. Tidal error and amplitude units are metres, and phase units are degrees.

Observation			Amplitude Error			Phase Error			Tidal Error		
Station name	**Ampl**	**Phase**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**
Cutler Farris	0.422	60.92	+0.011	-0.005	-	0	+5.21	-	0.008	0.027	-
Eastport	0.539	66.37	+0.048	-0.010	-	+5.93	+8.26	-	0.053	0.055	-
Saint John	0.615	65.74	+0.026	-0.005	-0.004	+2.57	+3.71	-0.44	0.027	0.028	0.005
Parkers Cove	0.679	56.82	+0.025	-0.028	-	+6.18	+6.16	-	0.056	0.054	-

Table 8: S2 water level tidal parameters for 2017. Tidal error and amplitude units are metres, and phase units are degrees.

Observation			Amplitude Error			Phase Error			Tidal Error		
Station name	**Ampl**	**Phase**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**
Cutler Farris	0.314	131.34	+0.001	+0.015	-	-6.29	+5.51	-	0.024	0.024	-
Eastport	0.419	138.70	+0.026	+0.012	-	-0.88	+8.37	-	0.019	0.045	-
Saint John	0.472	137.97	+0.018	+0.036	-0.001	-4.09	+3.75	-0.46	0.028	0.034	0.003
Parkers Cove	0.537	131.19	+0.009	+0.011	-	-2.98	+3.59	-	0.021	0.025	-

Table 9: K1 water level tidal parameters for 2017. Tidal error and amplitude units are metres, and phase units are degrees.

Observation			Amplitude Error			Phase Error			Tidal Error		
Station name	**Ampl**	**Phase**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**
Cutler Farris	0.148	193.89	+0.014	-0.003	-	-0.53	-4.35	-	0.010	0.008	-
Eastport	0.158	196.96	+0.018	-0.004	-	+0.16	-3.70	-	0.013	0.008	-
Saint John	0.160	194.59	+0.016	-0.002	-0.003	+0.58	-5.01	+0.78	0.011	0.010	0.003
Parkers Cove	0.165	190.27	+0.014	-0.005	-	+1.64	-4.17	-	0.011	0.009	-

Table 10: O1 water level tidal parameters for 2017. Tidal error and amplitude units are metres, and phase units are degrees.

Observation			Amplitude Error			Phase Error			Tidal Error		
Station name	**Ampl**	**Phase**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**
Cutler Farris	0.112	173.04	+0.010	-0.002	-	-3.36	-6.23	-	0.009	0.009	-
Eastport	0.116	176.23	+0.016	-0.001	-	-3.64	-5.93	-	0.013	0.008	-
Saint John	0.117	172.06	+0.014	+0.001	+0.001	-1.25	-5.63	+0.67	0.010	0.008	0.001
Parkers Cove	0.123	171.00	+0.010	-0.004	-	-2.96	-7.47	-	0.008	0.012	-

Table 11: L2 water level tidal parameters for 2017. Tidal error and amplitude units are metres, and phase units are degrees.

Observation			Amplitude Error			Phase Error			Tidal Error		
Station name	**Ampl**	**Phase**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**
Cutler Farris	0.112	136.79	-0.051	-0.006	-	+0.25	-13.36	-	0.036	0.018	-
Eastport	0.149	137.25	-0.063	-0.006	-	-8.81	-16.57	-	0.046	0.030	-
Saint John	0.182	137.47	-0.081	-0.018	-0.011	-7.81	-18.34	+0.62	0.059	0.041	0.008
Parkers Cove	0.194	128.84	-0.081	-0.019	-	-9.15	-17.71	-	0.060	0.042	-

Table 12: NU2 water level tidal parameters for 2017. Tidal error and amplitude units are metres, and phase units are degrees.

Observation			Amplitude Error			Phase Error			Tidal Error		
Station name	**Ampl**	**Phase**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**
Cutler Farris	0.098	68.80	-0.080	+0.006	-	+17.64	-3.60	-	0.057	0.006	-
Eastport	0.131	73.08	-0.105	+0.003	-	+7.97	-2.62	-	0.075	0.005	-
Saint John	0.148	73.42	-0.120	+0.005	+0.013	+8.96	-7.59	-0.94	0.085	0.015	0.010
Parkers Cove	0.162	63.61	-0.131	0	-	+9.54	-4.54	-	0.093	0.009	-

Table 13: K2 water level tidal parameters for 2017. Tidal error and amplitude units are metres, and phase units are degrees.

Observation			Amplitude Error			Phase Error			Tidal Error		
Station name	**Ampl**	**Phase**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**
Cutler Farris	0.088	130.50	-0.007	-0.004	-	-7.46	+5.19	-	0.009	0.006	-
Eastport	0.113	137.94	0	-0.005	-	+0.98	+7.59	-	0.001	0.011	-
Saint John	0.133	139.19	-0.010	-0.004	-0.002	-4.71	+1.55	-6.14	0.010	0.004	0.010
Parkers Cove	0.153	128.50	-0.015	-0.013	-	+1.10	+5.27	-	0.011	0.013	-

Table 14: 2N2 water level tidal parameters for 2017. Tidal error and amplitude units are metres, and phase units are degrees.

Observation			Amplitude Error			Phase Error			Tidal Error		
Station name	**Ampl**	**Phase**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**
Cutler Farris	0.060	33.98	-0.057	+0.013	-	+36.11	+22.05	-	0.041	0.020	-
Eastport	0.075	37.44	-0.073	+0.013	-	+4.85	+25.04	-	0.052	0.027	-
Saint John	0.086	36.55	-0.084	+0.015	-0.011	-5.01	+21.05	-5.02	0.059	0.026	0.009
Parkers Cove	0.091	25.93	-0.089	+0.016	-	-19.59	+25.19	-	0.063	0.032	-

Table 15: λ2 water level tidal parameters for 2017. Tidal error and amplitude units are metres, and phase units are degrees.

Observation			Amplitude Error			Phase Error			Tidal Error		
Station name	**Ampl**	**Phase**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**
Cutler Farris	0.035	115.00	-0.022	-0.002	-	-33.93	+0.58	-	0.018	0.002	-
Eastport	0.050	115.44	-0.029	+0.001	-	-38.93	-3.10	-	0.025	0.002	-
Saint John	0.059	114.80	-0.037	-0.001	-0.001	-36.43	-3.11	+7.02	0.030	0.002	0.005
Parkers Cove	0.065	107.80	-0.039	-0.003	-	-38.47	-4.68	-	0.033	0.004	-

Table 16: P1 water level tidal parameters for 2017. Tidal error and amplitude units are metres, and phase units are degrees.

Observation			Amplitude Error			Phase Error			Tidal Error		
Station name	**Ampl**	**Phase**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**
Cutler Farris	0.050	192.81	+0.002	-0.005	-	+0.76	+127.39	-	0.001	0.060	-
Eastport	0.054	195.96	+0.003	-0.007	-	+1.73	+127.68	-	0.002	0.064	-
Saint John	0.053	191.42	+0.004	-0.007	-0.003	+2.82	+128.54	+3.60	0.003	0.064	0.003
Parkers Cove	0.059	191.97	-0.001	-0.011	-	-0.87	+126.34	-	0.001	0.067	-

Table 17: M2 water level tidal parameters for Saint John for each year of the hindcast period. Tidal error and amplitude units are metres, and phase units are degrees.

Observation			Amplitude Error			Phase Error			Tidal Error		
Year	**Ampl**	**Phase**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**
2016	3.073	98.10	+0.240	-0.015	+0.001	-0.19	-1.02	-0.60	0.170	0.040	0.033
2017	3.055	97.92	+0.249	+0.023	+0.037	+0.01	-0.72	-0.44	0.176	0.032	0.031
2018	3.079	97.79	+0.231	+0.009	+0.022	+0.47	-0.20	+0.07	0.164	0.010	0.016
2019	3.026	99.03	+0.278	+0.054	+0.068	-0.94	-1.59	-1.20	0.200	0.071	0.066

Observation			Amplitude Error			Phase Error			Tidal Error		
Year	Ampl	Phase	CIOPS-E	Fundy500	SJ100	CIOPS-E	Fundy500	SJ100	CIOPS-E	Fundy500	SJ100
2016	3.073	98.10	+0.240	-0.015	+0.001	-0.19	-1.02	-0.60	0.170	0.040	0.033
2020	3.046	98.44	+0.264	+0.033	+0.048	-0.55	-1.08	-0.63	0.188	0.047	0.042
2021	3.067	97.95	+0.284	+0.060	+0.074	+0.13	-0.38	+0.01	0.201	0.045	0.052

Table 18: Total water level statistics for 2018.

Score:	Bias (m)			CRMSE (m)			χ^2			pearson		
Model:	CIOPS-E	Fundy500	SJ100	CIOPS-E	Fundy500	SJ100	CIOPS-E	Fundy500	SJ100	CIOPS-E	Fundy500	SJ100
Cutler Farris	0.026	0.009	-	0.15	0.123	-	0.009	0.006	-	0.997	0.997	-
Eastport	-0.099	-0.1	-	0.314	0.194	-	0.024	0.009	-	0.995	0.995	-
Saint John	-0.033	-0.033	-0.022	0.245	0.145	0.097	0.011	0.004	0.002	0.997	0.998	0.999

Table 19: Tidal water level statistics for 2018.

Score:	CRMSE (m)			χ^2			pearson		
Model:	CIOPS-E	Fundy500	SJ100	CIOPS-E	Fundy500	SJ100	CIOPS-E	Fundy500	SJ100
Cutler Farris	0.137	0.106	-	0.008	0.005	-	0.997	0.998	-
Eastport	0.3	0.175	-	0.022	0.008	-	0.996	0.996	-
Saint John	0.227	0.118	0.06	0.01	0.003	0.001	0.998	0.999	1

Table 20: Non-tidal water level statistics for 2018.

Score:	CRMSE (m)			χ^2			pearson		
Model:	CIOPS-E	Fundy500	SJ100	CIOPS-E	Fundy500	SJ100	CIOPS-E	Fundy500	SJ100
Cutler Farris	0.069	0.063	-	0.246	0.203	-	0.877	0.898	-
Eastport	0.099	0.086	-	0.445	0.329	-	0.791	0.832	-
Saint John	0.107	0.084	0.079	0.572	0.353	0.315	0.728	0.827	0.843

Table 21: Statistics for storm surge, averaged over up to 11 storms. Number of storms available for each station is indicated in brackets next to the station name.

Score:	CRMSE (m)			χ^2			pearson		
Model:	CIOPS-E	Fundy500	SJ100	CIOPS-E	Fundy500	SJ100	CIOPS-E	Fundy500	SJ100
Cutler Farris (11)	0.046	0.046	-	0.131	0.125	-	0.934	0.941	-

Score:	CRMSE (m)			χ^2			pearson		
Model:	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**	**CIOPS-E**	**Fundy500**	**SJ100**
Eastport (11)	0.061	0.053	-	0.197	0.138	-	0.907	0.934	-
Saint John (11)	0.059	0.059	0.056	0.205	0.192	0.178	0.893	0.907	0.911
Parkers Cove (1)	0.056	0.055	-	0.16	0.158	-	0.932	0.943	-

Table 22: SST statistics for entire hindcast period (2016-2021). Note that the Fundy FORCE scores may not be robust, as there are cold spikes due to air temperature being measured rather than water temperature (see Figure 59). Units for bias and CRMSE are °C, and γ^2 and Pearson are unitless.

Station Name	Score	CIOPS-E	Fundy500	SJ100
44490 – West Bay of Fundy	**bias**	0.122	0.481	-
	CRMSE	0.738	0.87	-
	γ^2	0.049	0.068	-
	pearson	0.986	0.986	-
Cutler Farris	**bias**	0.89	0.752	-
	CRMSE	1.238	1.04	-
	γ^2	0.137	0.097	-
	pearson	0.963	0.969	-
Eastport	**bias**	1.794	0.884	-
	CRMSE	2.344	1.041	-
	γ^2	0.584	0.115	-
	pearson	0.936	0.98	-
FundyFORCE	**bias**	0.391	0.438	-
	CRMSE	0.89	1.071	-
	γ^2	0.031	0.045	-
	pearson	0.986	0.983	-
SmartAtlantic Saint John	**bias**	-0.221	-0.064	-0.33
	CRMSE	1.268	1.356	1.161
	γ^2	0.081	0.092	0.068
	pearson	0.962	0.953	0.966

Table 23: Statistics for moored CTDs. Units for bias and CRMSE are °C and psu for temperature and salinity respectively; γ^2 and Pearson are unitless.

Station Name	Variable: Model:	Temperature CIOPS-E	Temperature Fundy500	Temperature SJ100	Salinity CIOPS-E	Salinity Fundy500	Salinity SJ100
2090_183 12019-04	**bias**	0.517	0.563	-	-	-	-
	CRMSE	0.904	0.872	-	-	-	-
	γ^2	0.241	0.224	-	-	-	-
	pearson	0.964	0.969	-	-	-	-
011 Courtenay Bay Intertidal	**bias**	-	-0.509	-0.186	-	-2.264	-3.348
	CRMSE	-	2.384	1.861	-	13.084	9.991
	γ^2	-	1.284	0.782	-	2.243	1.308
	pearson	-	0.502	0.588	-	-0.19	-0.019
013 SJR Narrows	**bias**	-	-0.045	-0.037	-	3.599	-6.191
	CRMSE	-	0.454	0.851	-	5.37	5.399
	γ^2	-	0.151	0.531	-	0.801	0.81
	pearson	-	0.954	0.954	-	0.569	0.552
050 Musquash	**bias**	-	0.628	0.49	-	2.25	2.264
	CRMSE	-	0.494	0.416	-	2.106	2.116

Station Name	Variable: Model:	Temperature CIOPS-E	Temperature Fundy500	Temperature SJ100	Salinity CIOPS-E	Salinity Fundy500	Salinity SJ100
East Headlands	γ^2	-	0.286	0.203	-	1.161	1.171
	pearson	-	0.944	0.947	-	-0.288	-0.388
059 Musquash East Headlands	bias	-	0.157	0.114	-	0.488	0.582
	CRMSE	-	0.508	0.478	-	0.846	0.838
	γ^2	-	0.053	0.047	-	0.75	0.737
	pearson	-	0.974	0.977	-	0.56	0.539
060 Musquash West Headlands	bias	-	0.174	-0.003	-	0.509	0.63
	CRMSE	-	0.51	0.602	-	0.856	0.818
	γ^2	-	0.071	0.099	-	0.66	0.603
	pearson	-	0.966	0.95	-	0.605	0.631

Table 24: Mean Molcard skill score for each drifter type and model, for 6 simulation lengths.

Drifter Type	Hour	Molcard Skill CIOPS-E	Molcard Skill Fundy500	Molcard Skill SJ100	Separation Distance (km) CIOPS-E	Separation Distance (km) Fundy500	Separation Distance (km) SJ100
CODE-Davis	4	0.56	0.57	0.59	2.47	2.44	2.36
	8	0.44	0.43	0.46	4.13	4.25	4.07
	12	0.22	0.2	0.21	5.33	5.85	5.53
	16	0.32	0.27	0.29	6.12	7.35	6.83
	20	0.33	0.28	0.3	7.11	8.78	7.93
	24	0.25	0.21	0.22	7.9	10.16	8.52
SABS Long	4	0.52	0.58	0.59	3.33	2.73	2.7
	8	0.43	0.46	0.47	5.35	4.68	4.59
	12	0.24	0.3	0.31	6.68	6.24	5.91
	16	0.33	0.36	0.37	7.68	7.49	7.27
	20	0.36	0.38	0.39	8.26	8.38	7.86
	24	0.28	0.31	0.3	9.46	9.82	8.75
SABS Disk	4	0.49	0.52	0.52	3.4	3.21	3.3
	8	0.4	0.4	0.4	5.42	5.52	5.67
	12	0.31	0.28	0.28	6.36	7.15	7.05
	16	0.33	0.31	0.32	7.34	8.47	8.19
	20	0.23	0.29	0.3	9.54	9.55	8.84
	24	0.19	0.25	0.26	11.02	11.16	10.18

12 FIGURES

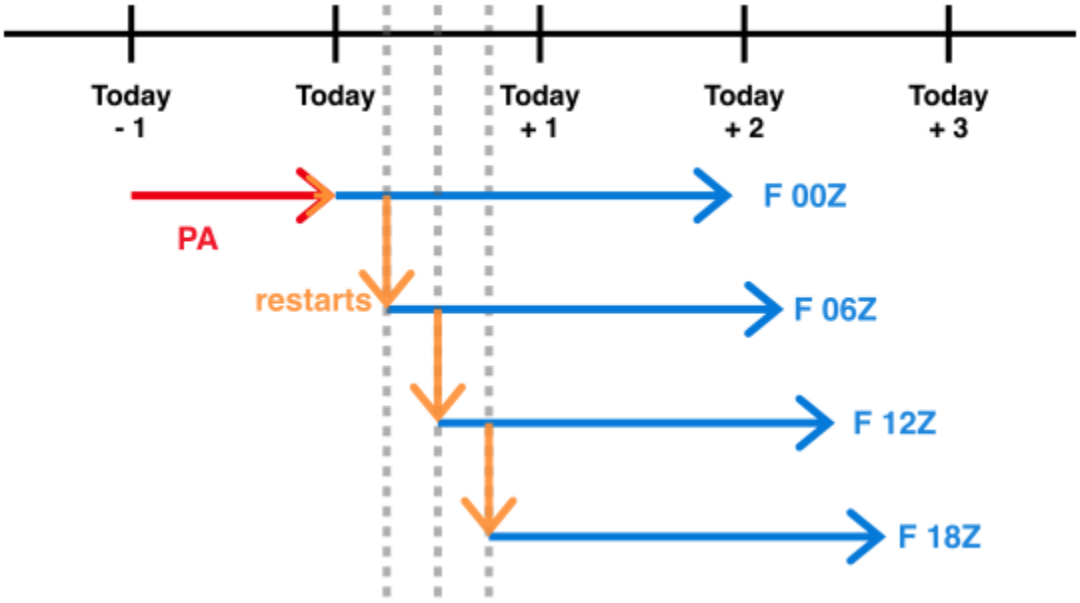

Figure 1: Schematic of one timestamp's set of pseudo-analysis (PA, in red) and forecast (in blue) runs. Grey dashed lines are spaced six hours apart, and orange arrows indicate where a restart file is generated and used to launch the subsequent step. The PA for today+1 will start with the same restart used to start today's 00Z forecast, and the pattern will repeat.

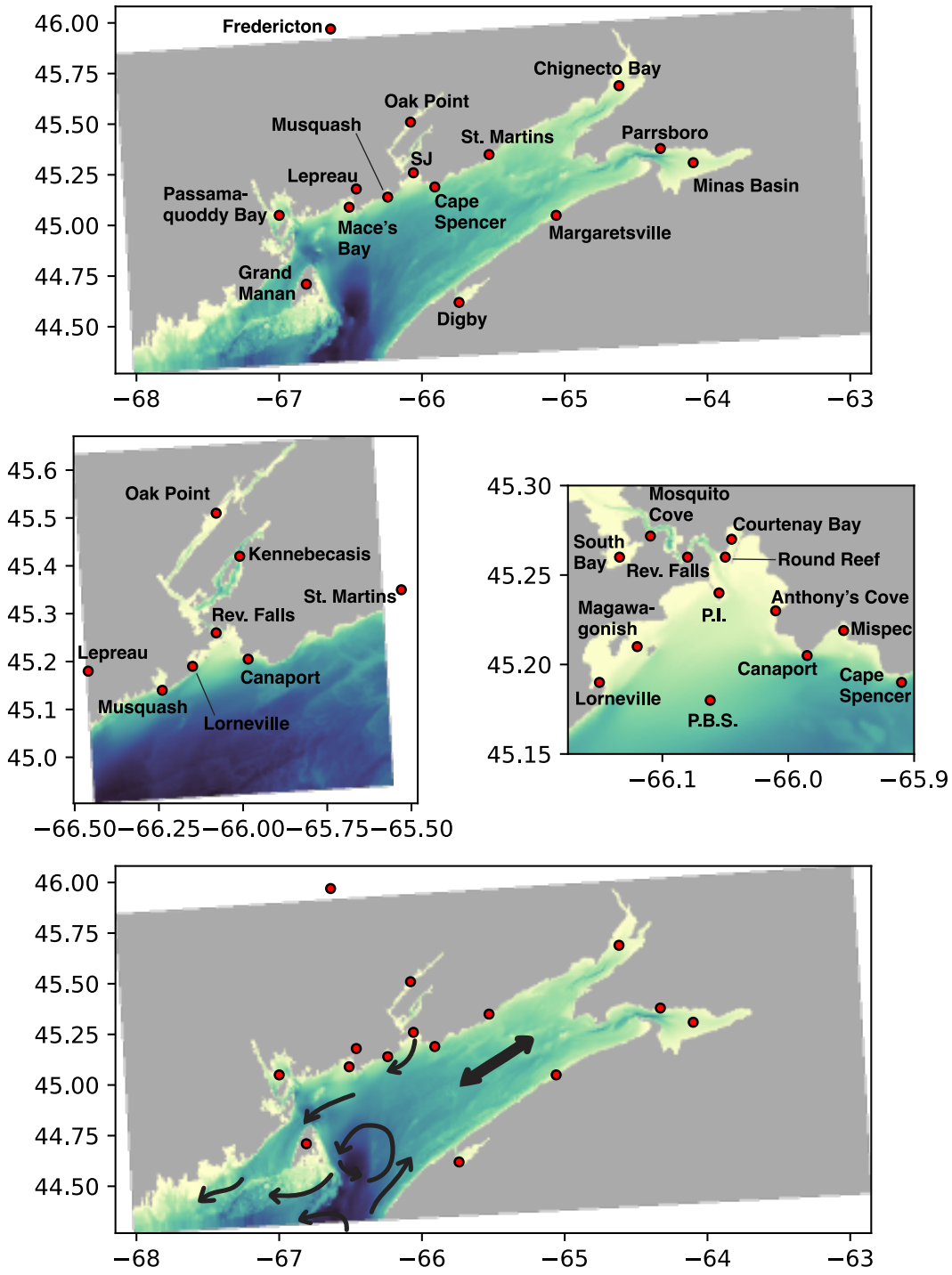

Figure 2: Map of Fundy500 (top) and SJ100 (middle left) domains with notable locations labelled. A zoomed in section of SJ100 around the port of Saint John (middle right) shows more local features in the harbour. The bottom panel is a schematic of general circulation patterns in the Bay of Fundy. P.I. is short for Partridge Island, and P.B.S. stands for Pilot Boarding Station

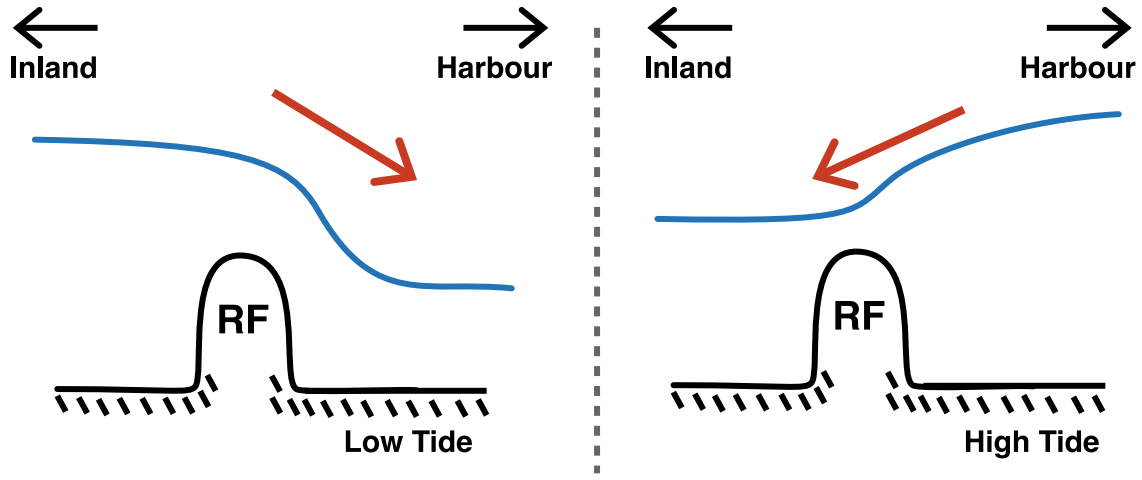

Figure 3: A schematic of water level at Reversing Falls (RF), with both low and high tide patterns shown. Water level is shown with a blue line, and the orange arrow indicates the direction of flow.

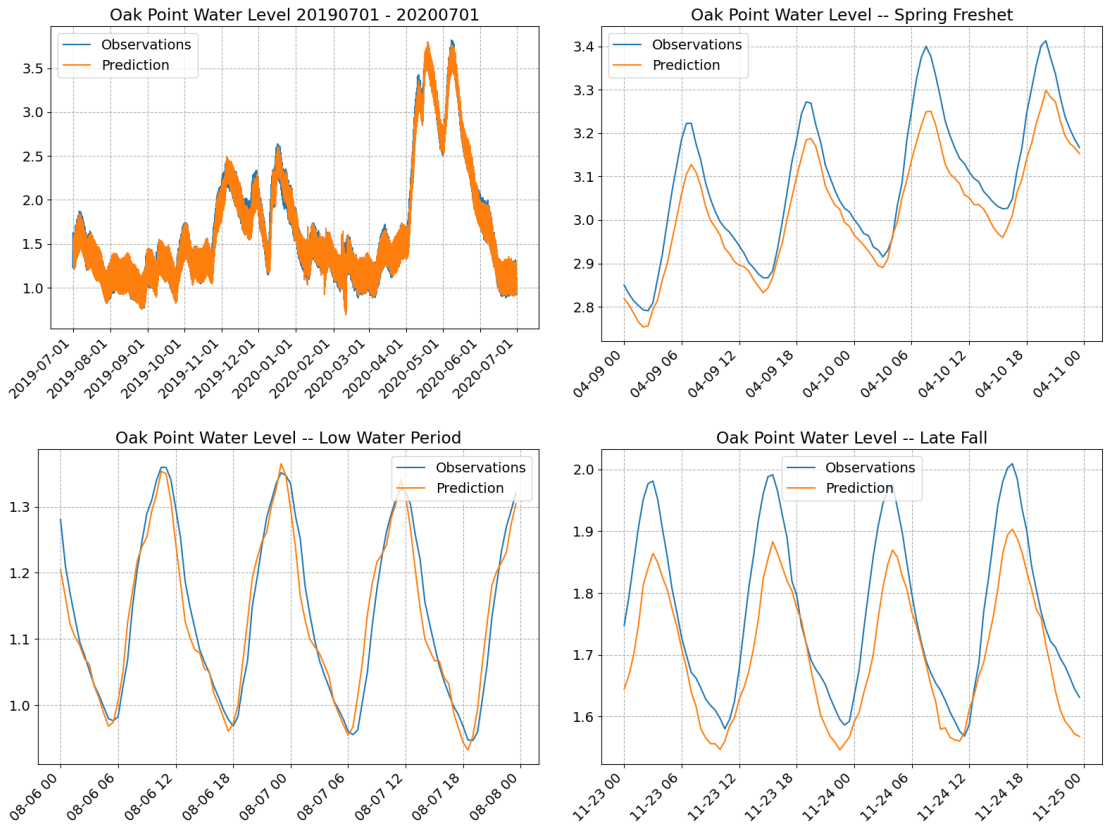

Figure 4: Water level observed and predicted by NS_TIDE at Oak Point. Top left shows the basis used for all years, and other three panels show short, two-day snippets of observations and NS_TIDE prediction at various river stages.

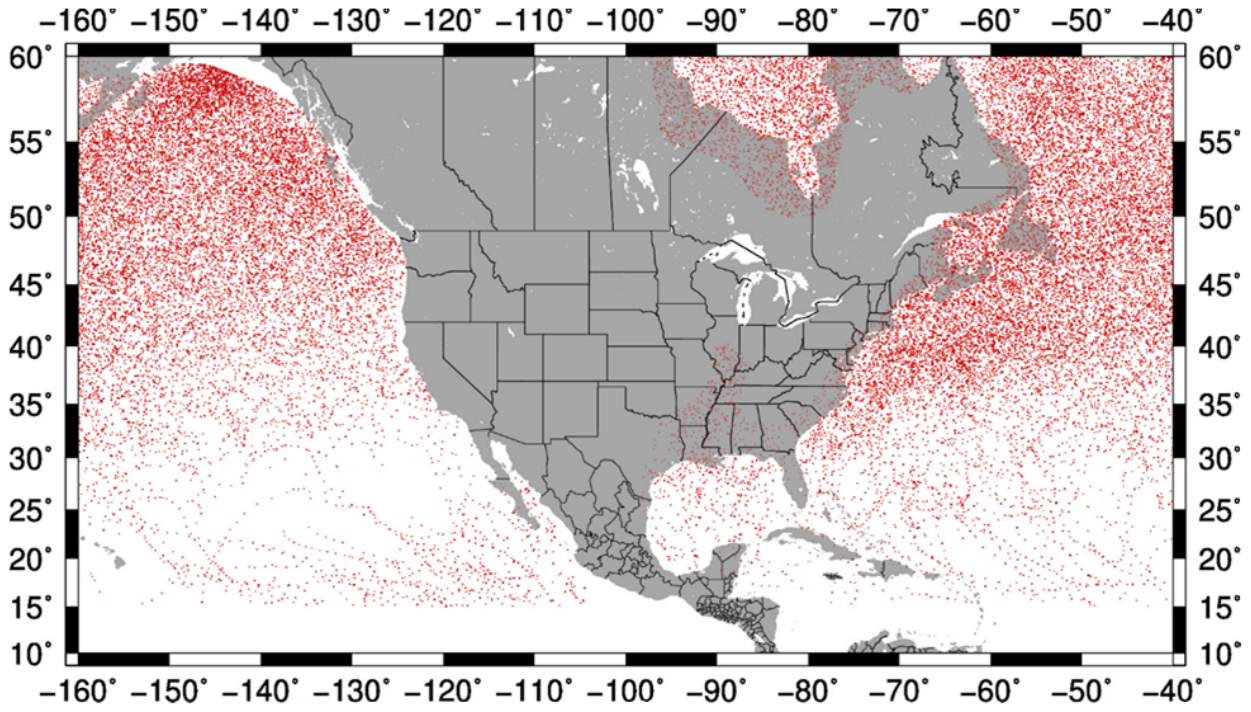

Figure 5. Cyclone locations every six hours from 2010-2021.

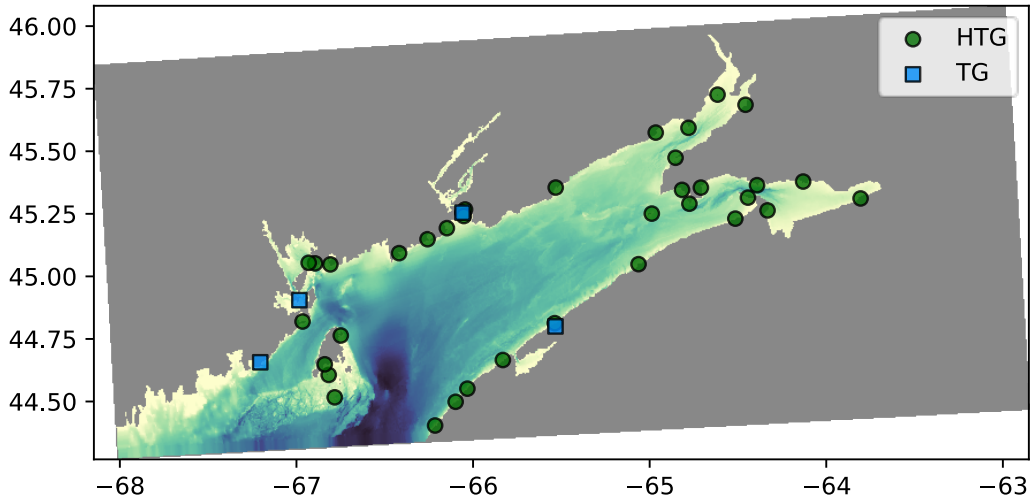

Figure 6: Tide gauges (blue squares) and tidal constituent stations (green circles) in the Bay of Fundy.

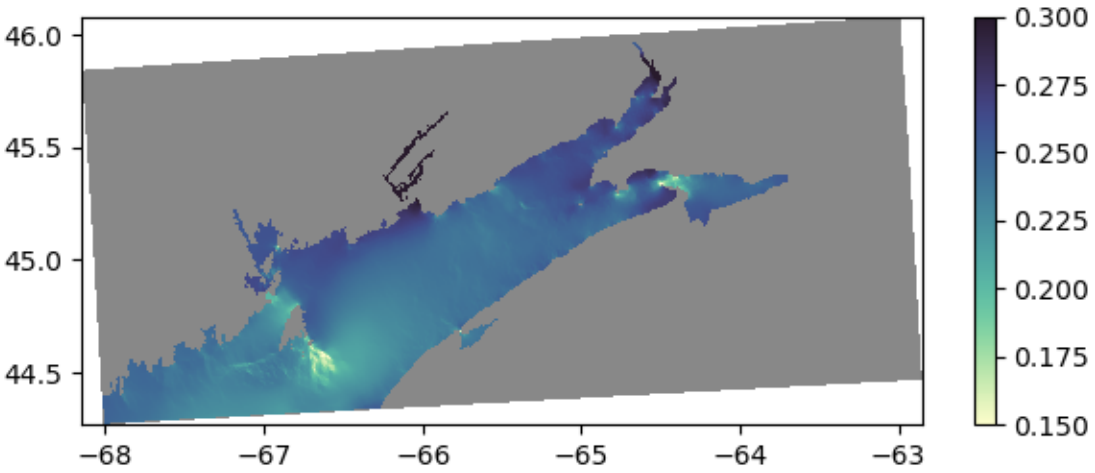

Figure 7: Hindcast mean SSH for Fundy500. Units are metres.

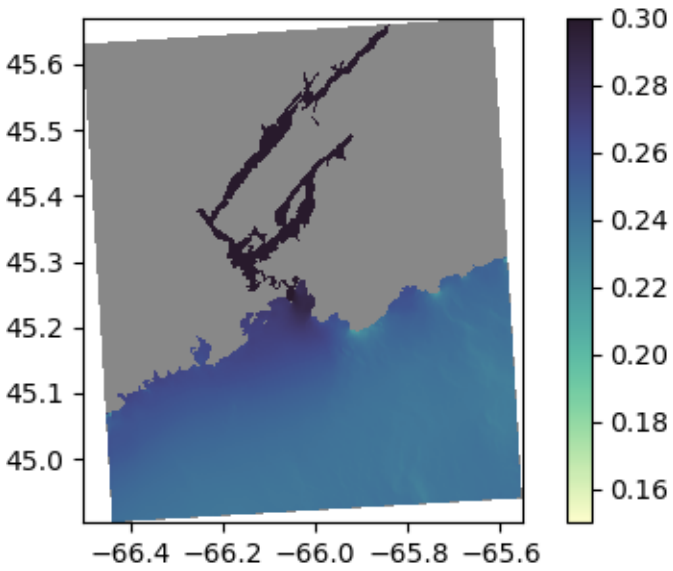

Figure 8: Hindcast mean SSH for SJ100. Color scale is the same as in Figure 7.

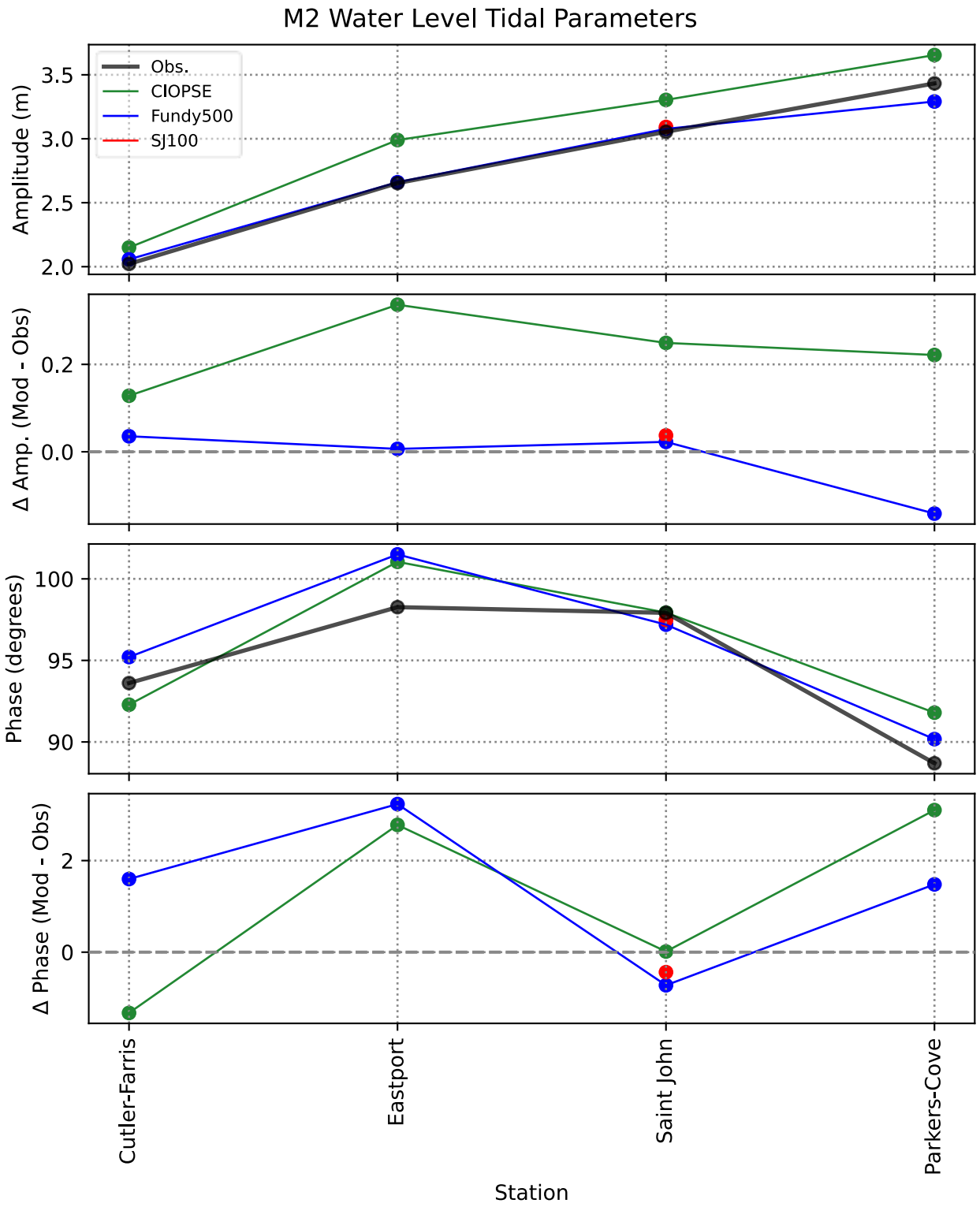

Figure 9: M2 tidal constituents for the four tide gauges available in the domain. Note that this shows data calculated over 2017 rather than 2018, so Parker's Cove can be included in the plot.

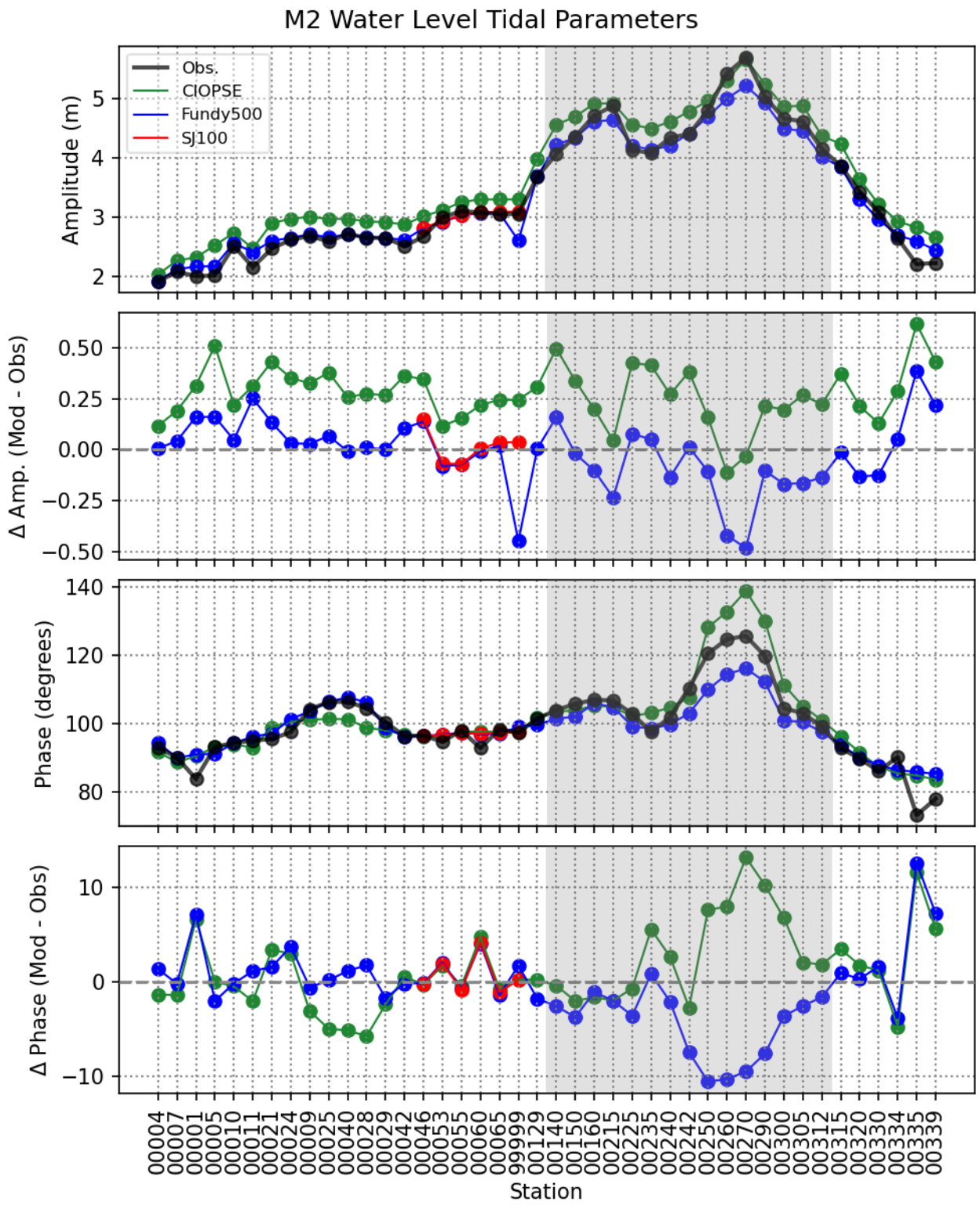

Figure 10: M2 constituent parameters for constituent-only stations around the Bay of Fundy. Cutler (00004) is on the northern coast of Maine, and stations progress clockwise around the Bay of Fundy. Stations along the Saint John River, along with some in Passamaquoddy Bay and in Upper Fundy, are omitted. The Upper Fundy region, with significant intertidal zones, is shaded.

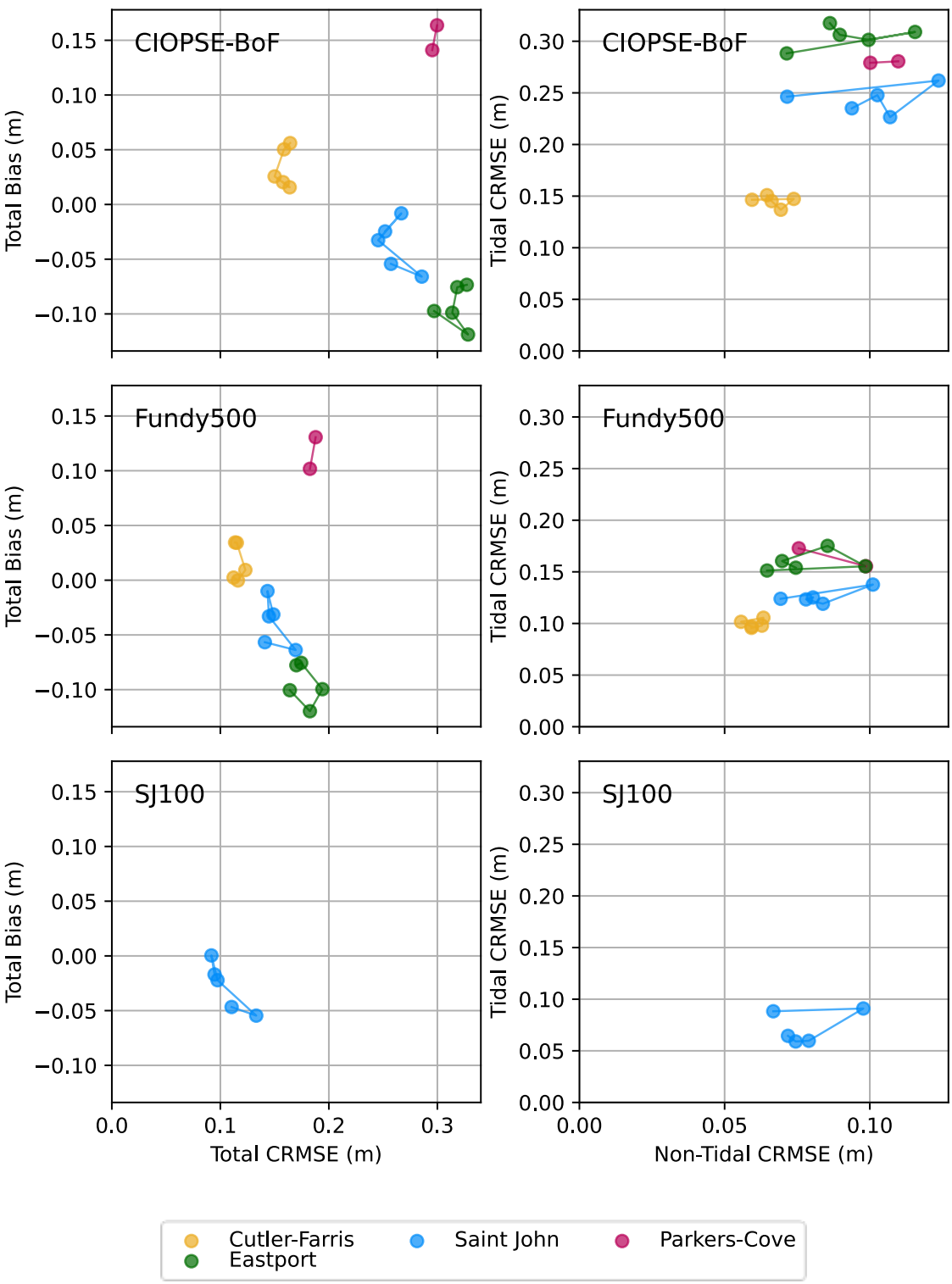

Figure 11: Key statistics for water level evaluation plotted yearly from 2016 to 2022.

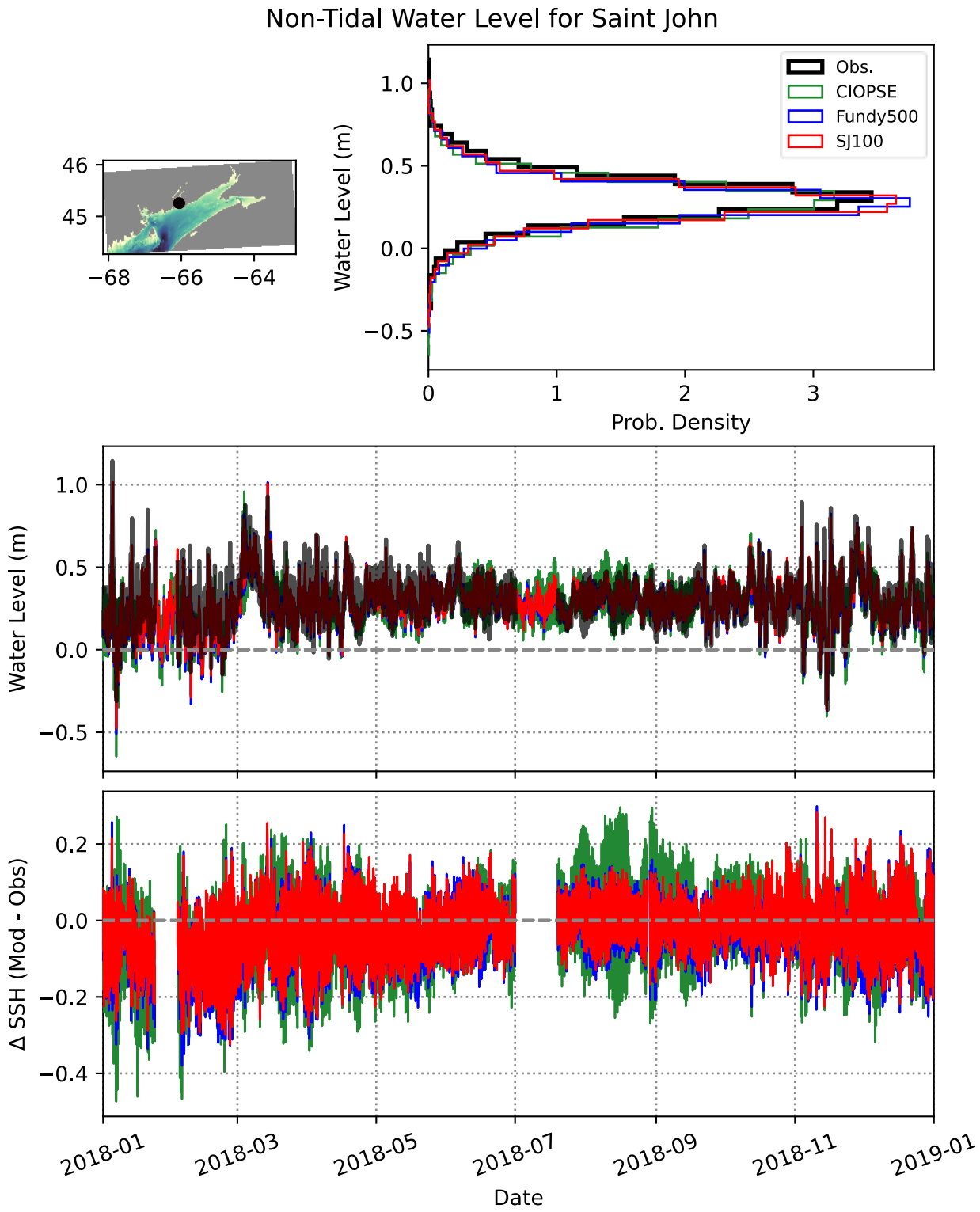

Figure 12: Non-tidal water level for Saint John for 2018. Note the increased non-tidal water level in the early part of the year, though the flood in April is more evident in the currents rather than water level in the harbour.

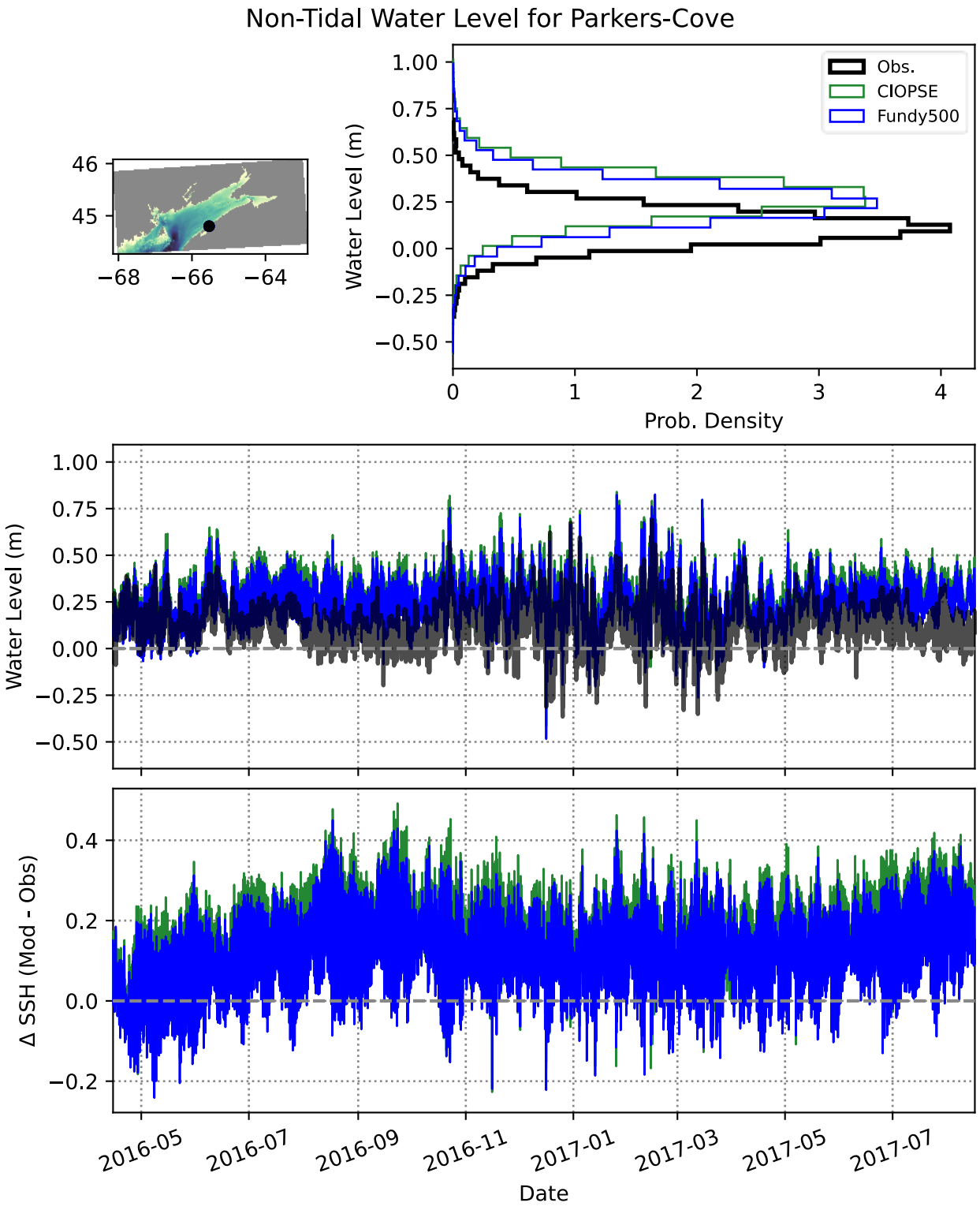

Figure 13: Non-tidal water level plotted for the entire time series available for Parker's Cove.

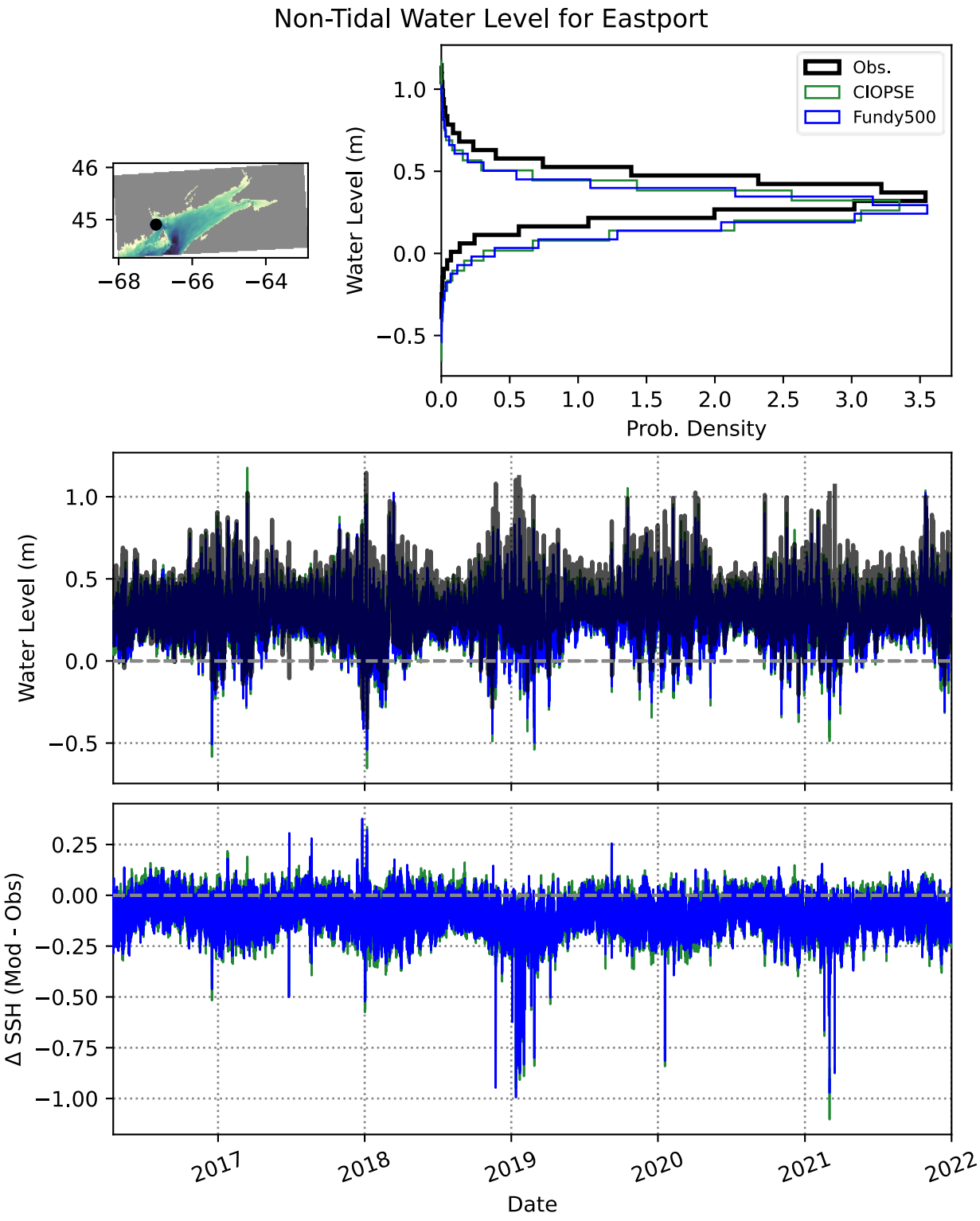

Figure 14: Non-tidal water level plotted for the entire hindcast period for Eastport.

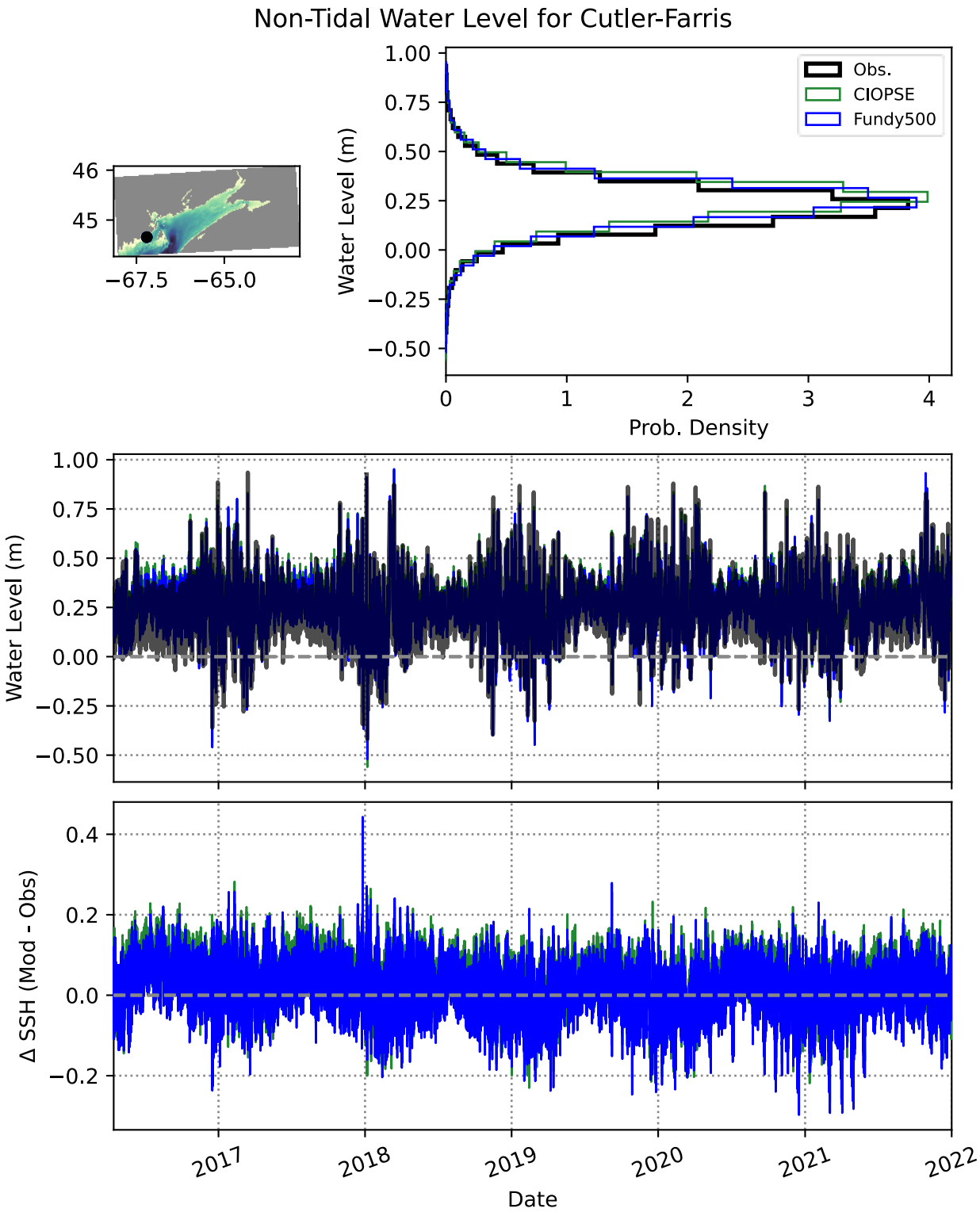

Figure 15: Non-tidal water level plotted for the entire hindcast period for Cutler Farris.

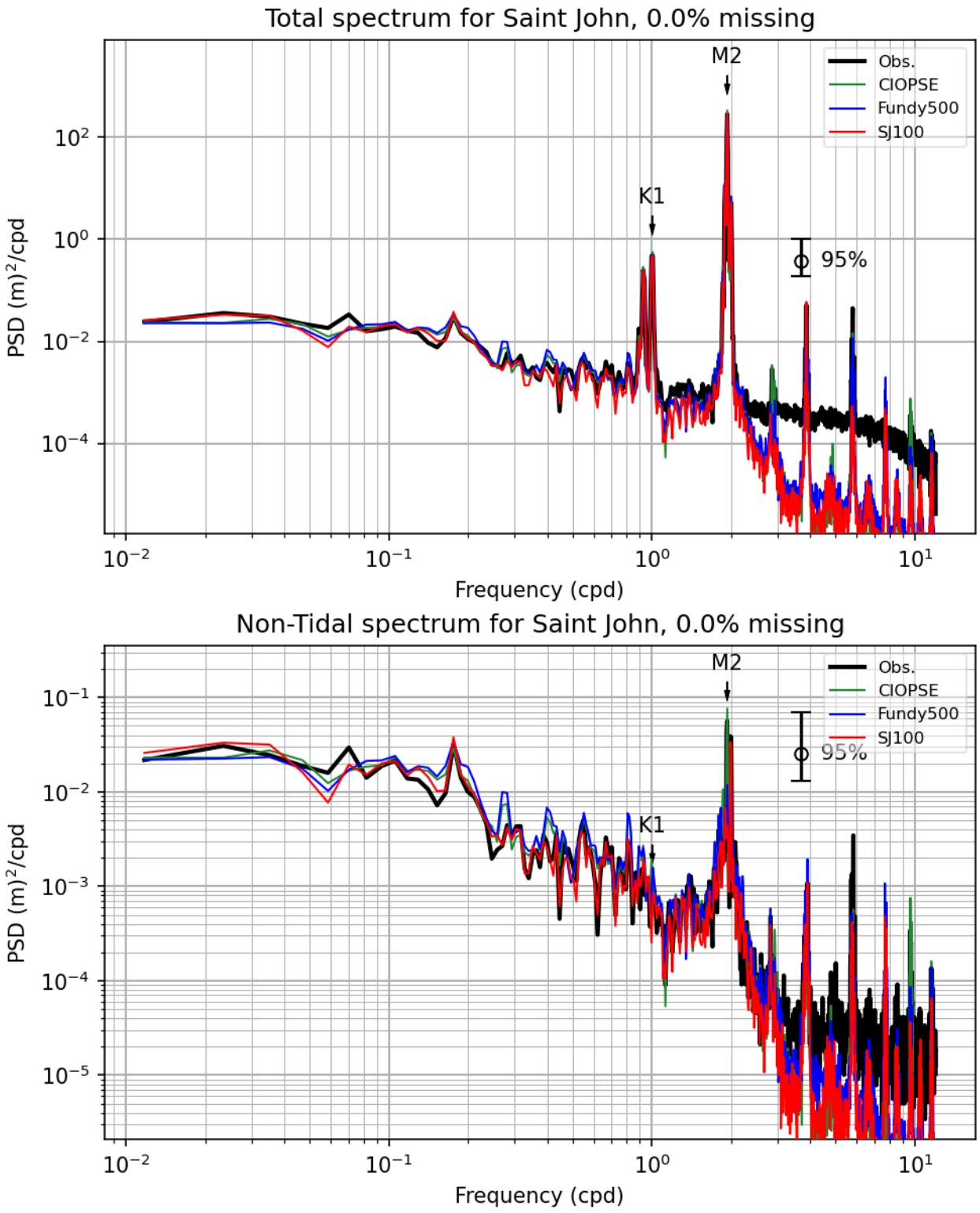

Figure 16: Frequency spectra for total (top) and non-tidal (bottom) water level at Saint John in 2016.

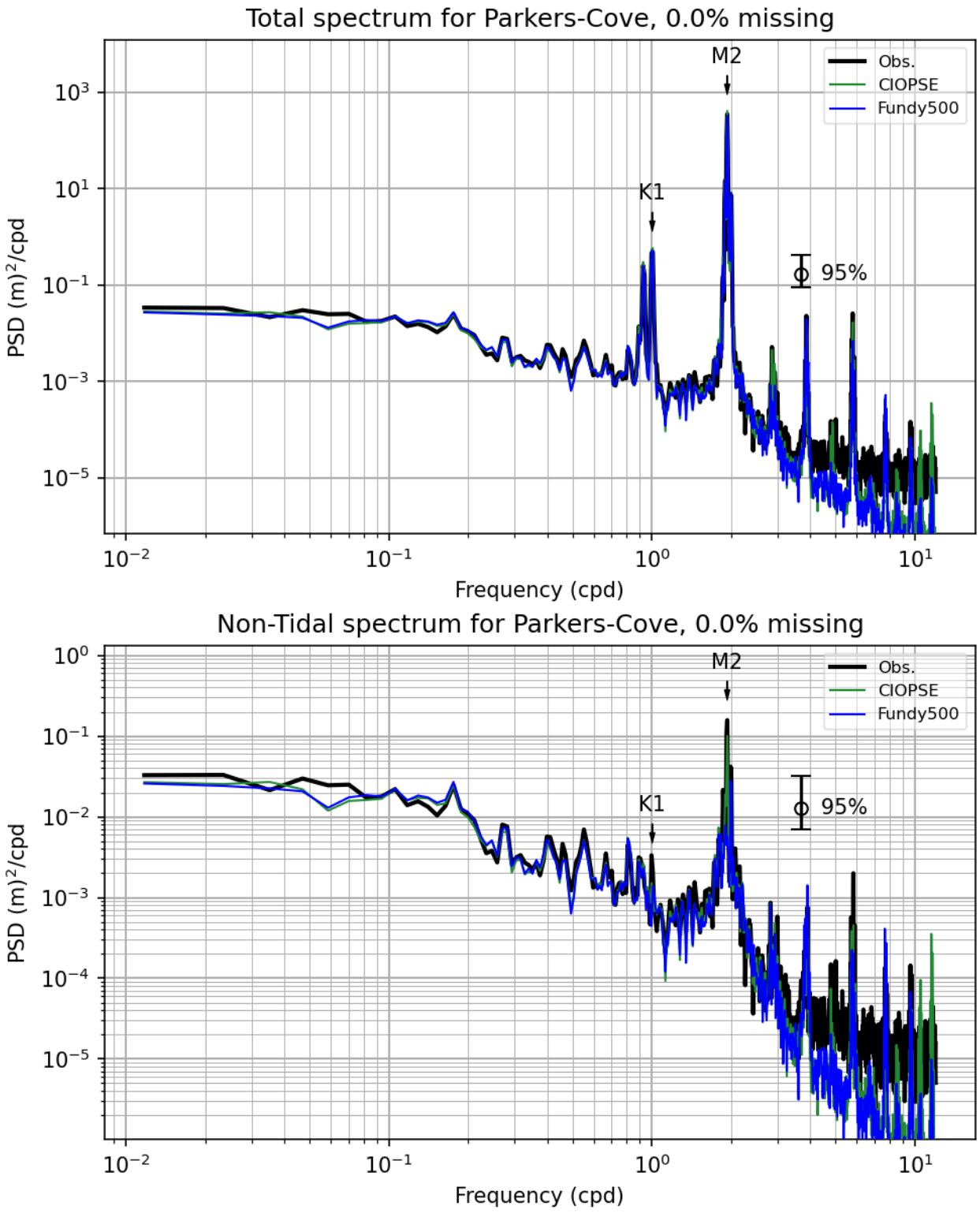

Figure 17: Frequency spectra for total (top) and non-tidal (bottom) water level at Parker's Cove in 2016.

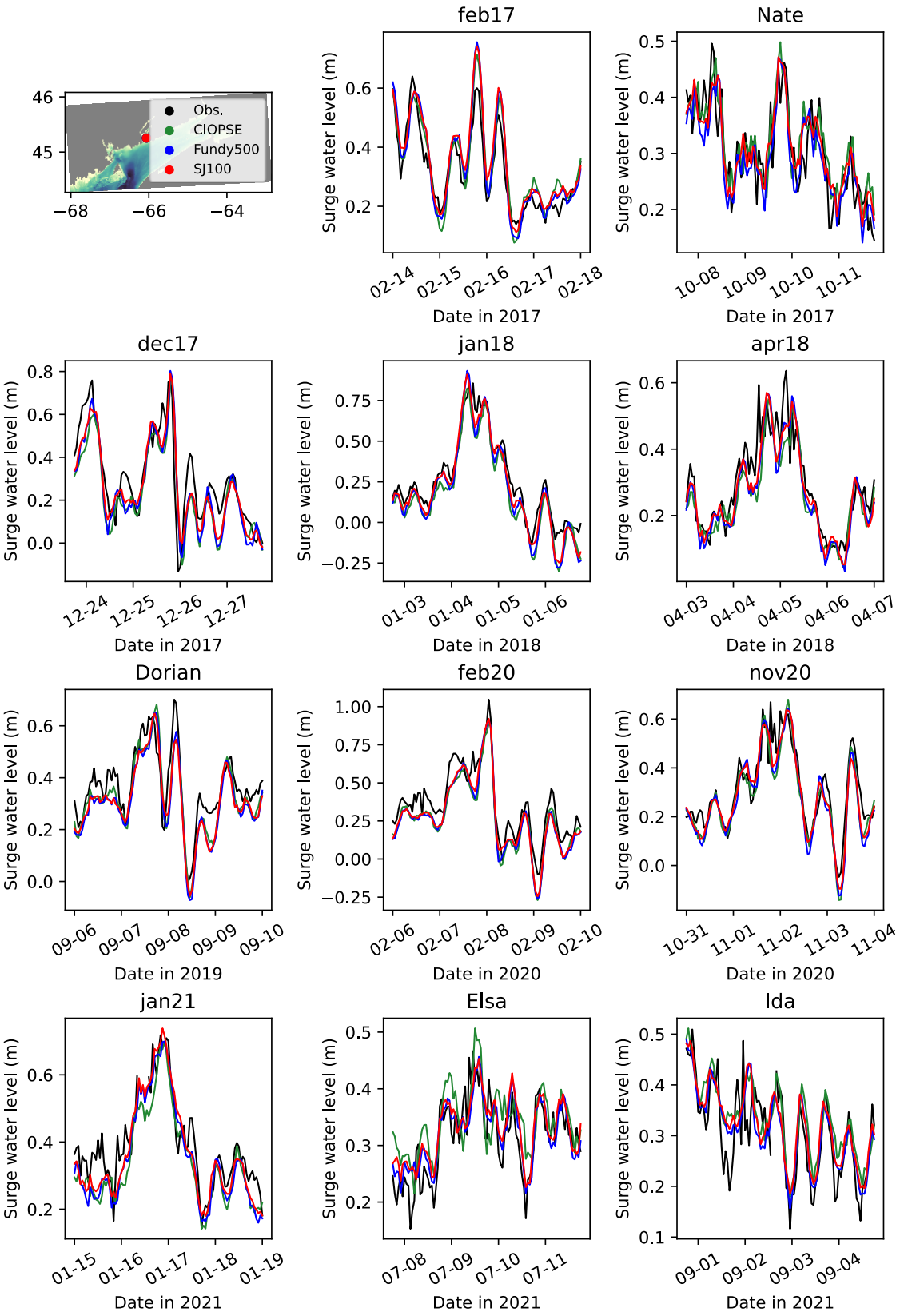

Figure 18: Storm surge at Saint John for 11 selected storms.

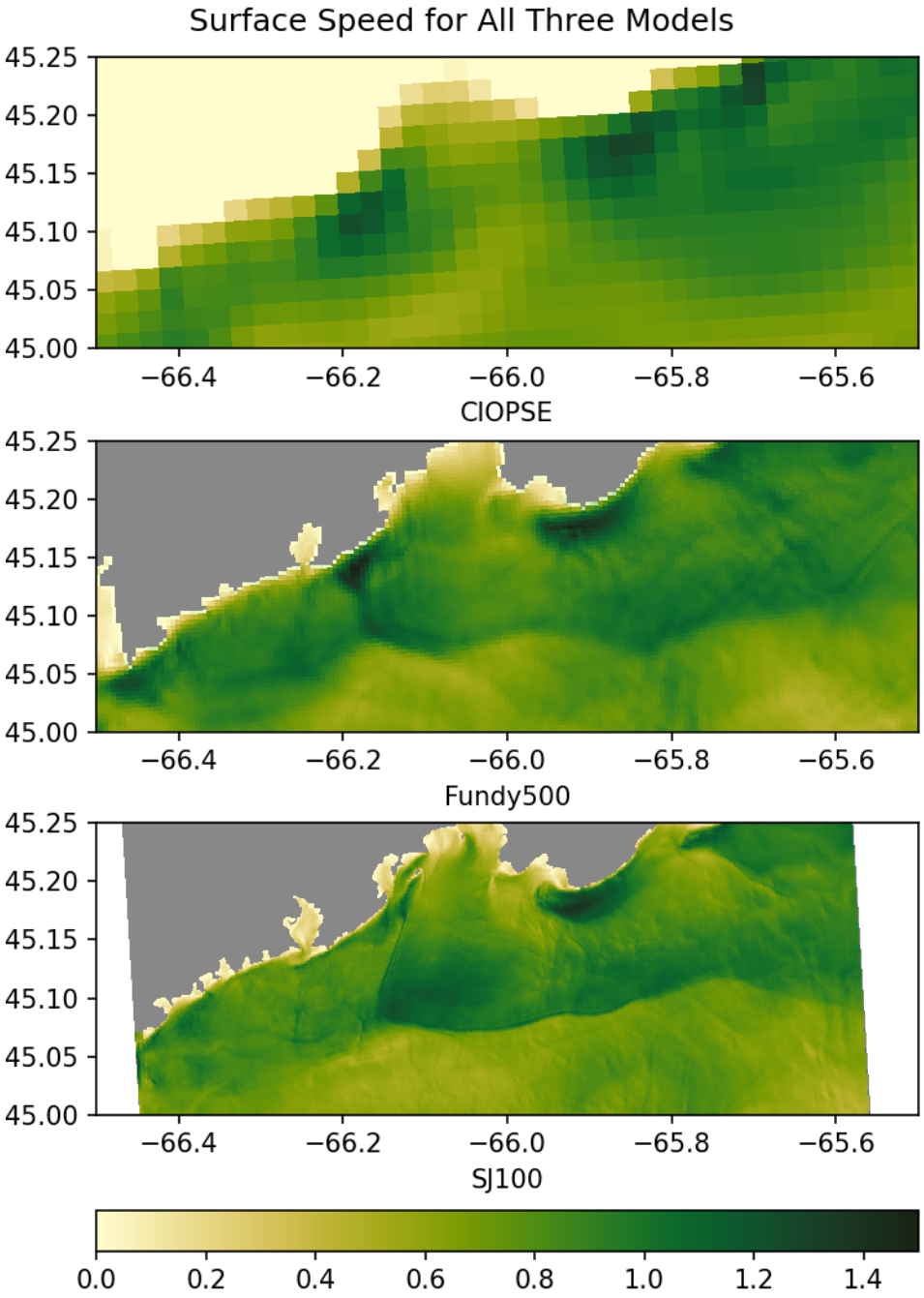

Figure 19: Surface speed at 2020-06-01 00:00 UTC for all three models. Note that features are sharper and more distinct in SJ100 than in the other two models. Units are m/s.

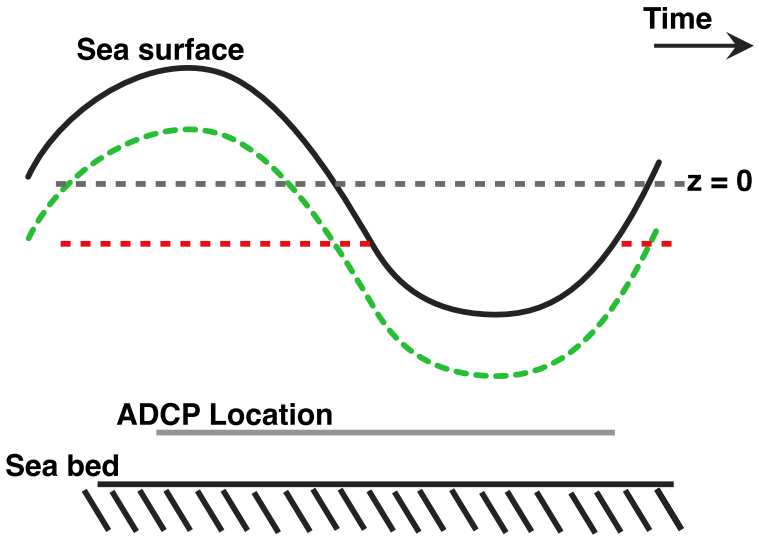

Figure 20: Schematic showing the importance of evaluating ADCP data relative to the variable surface (green dashed line) rather than the z=0 reference point (red dashed line). The red dashed line has a regular gap in data at low tide, which will alias any evaluation.

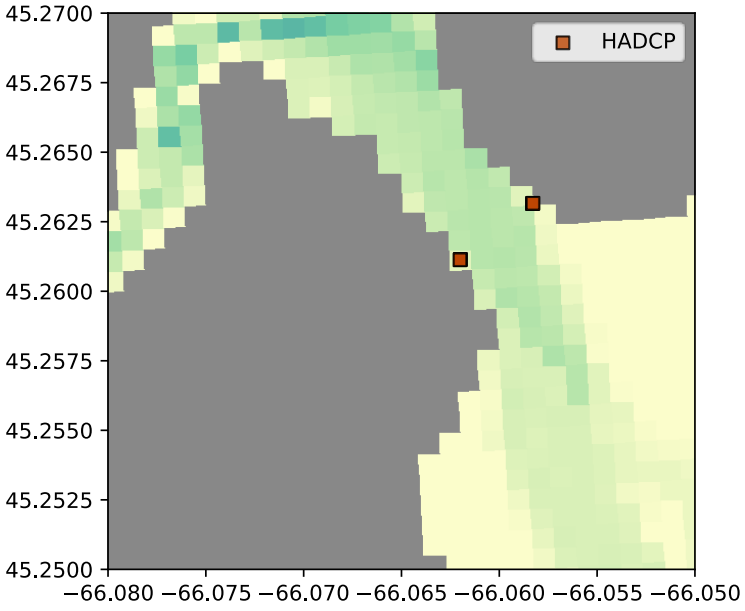

Figure 21: Locations of HADCP installations at the mouth of the Saint John River.

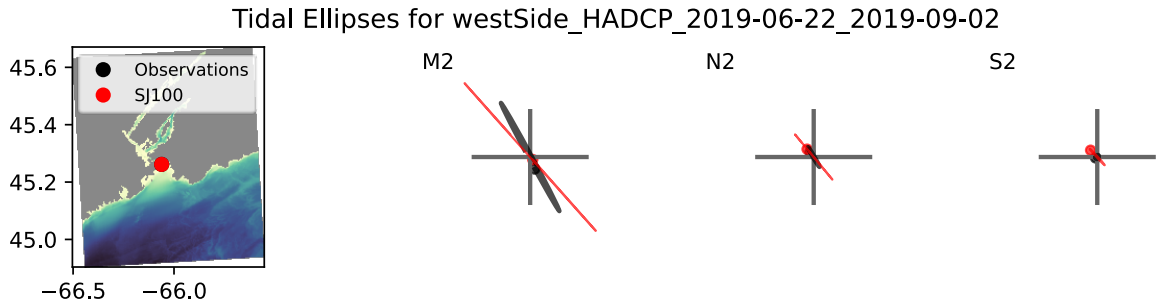

Figure 22: Tidal ellipses at the west side HADCP for mid-2019 deployment.

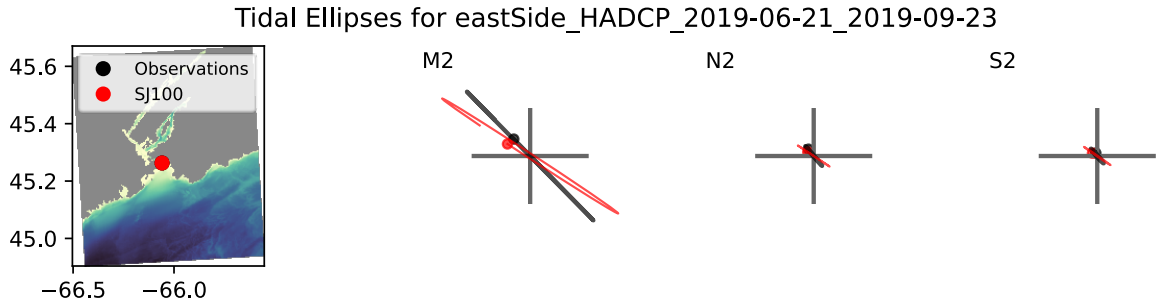

Figure 23: Tidal ellipses at the east side HADCP for mid-2019 deployment.

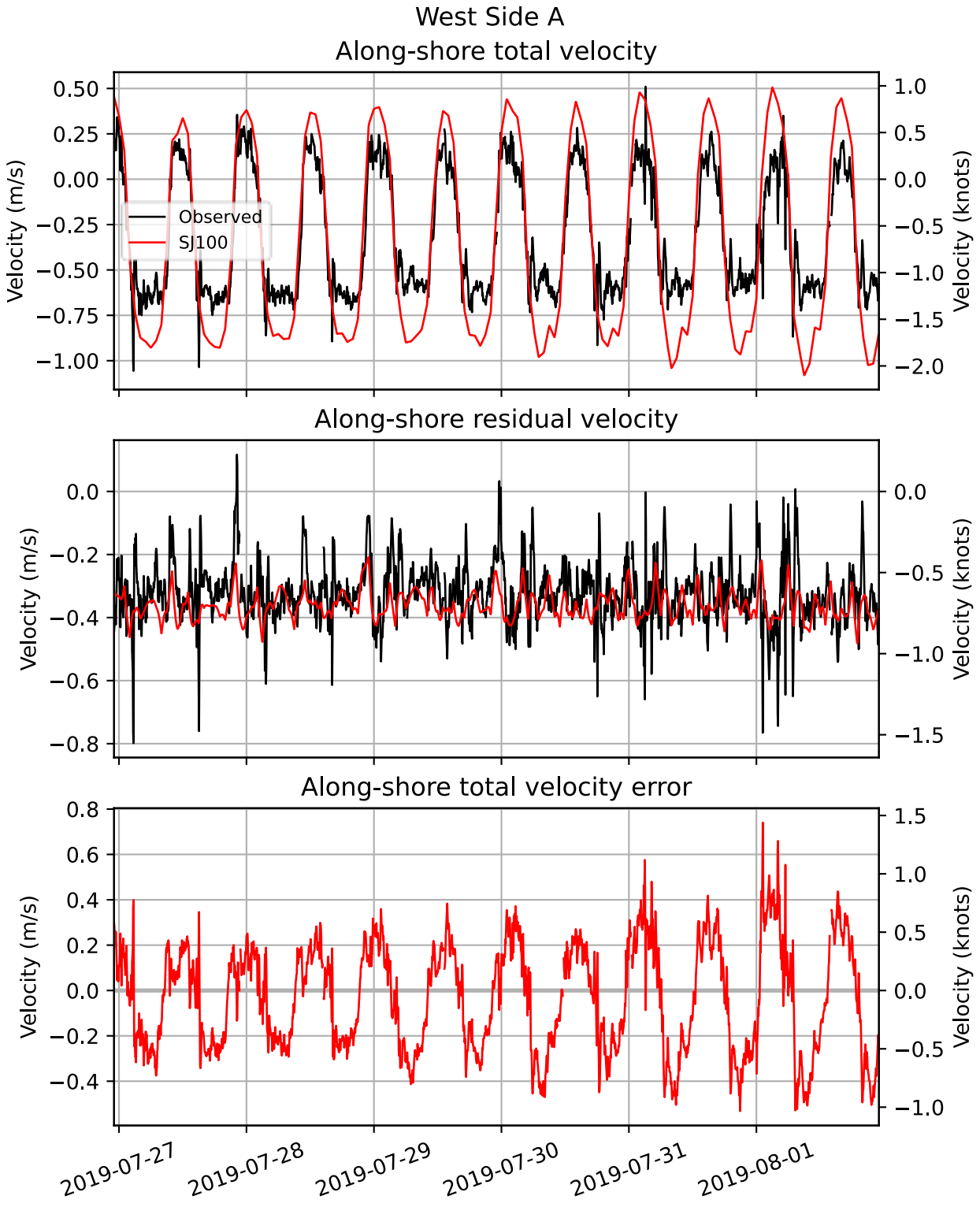

Figure 24: One week time series of along-shore velocities (total and non-tidal) at the western HADCP for mid-2019 deployment.

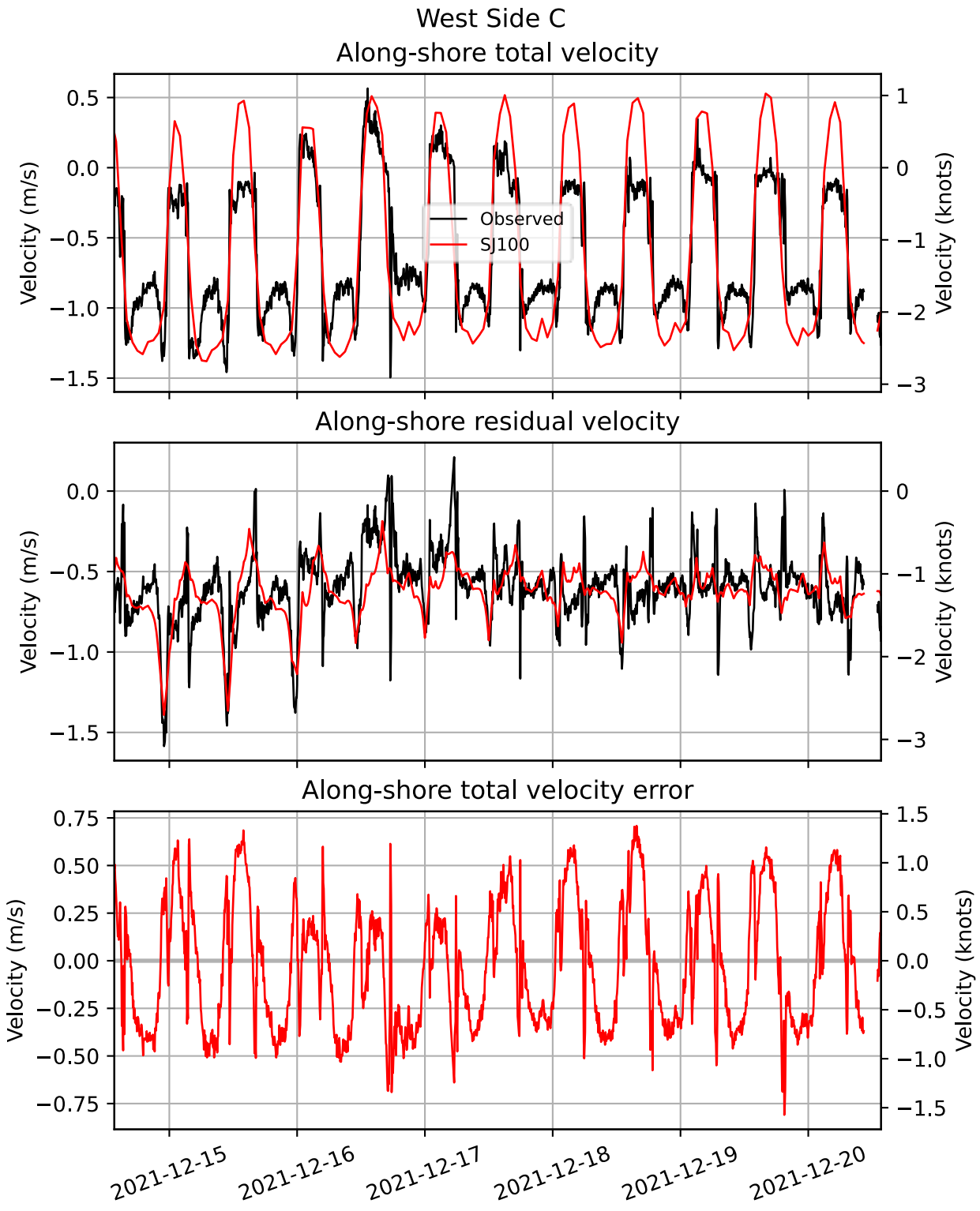

Figure 25: One-week timeseries of total and non-tidal along shore velocities at the western HADCP in December 2021.

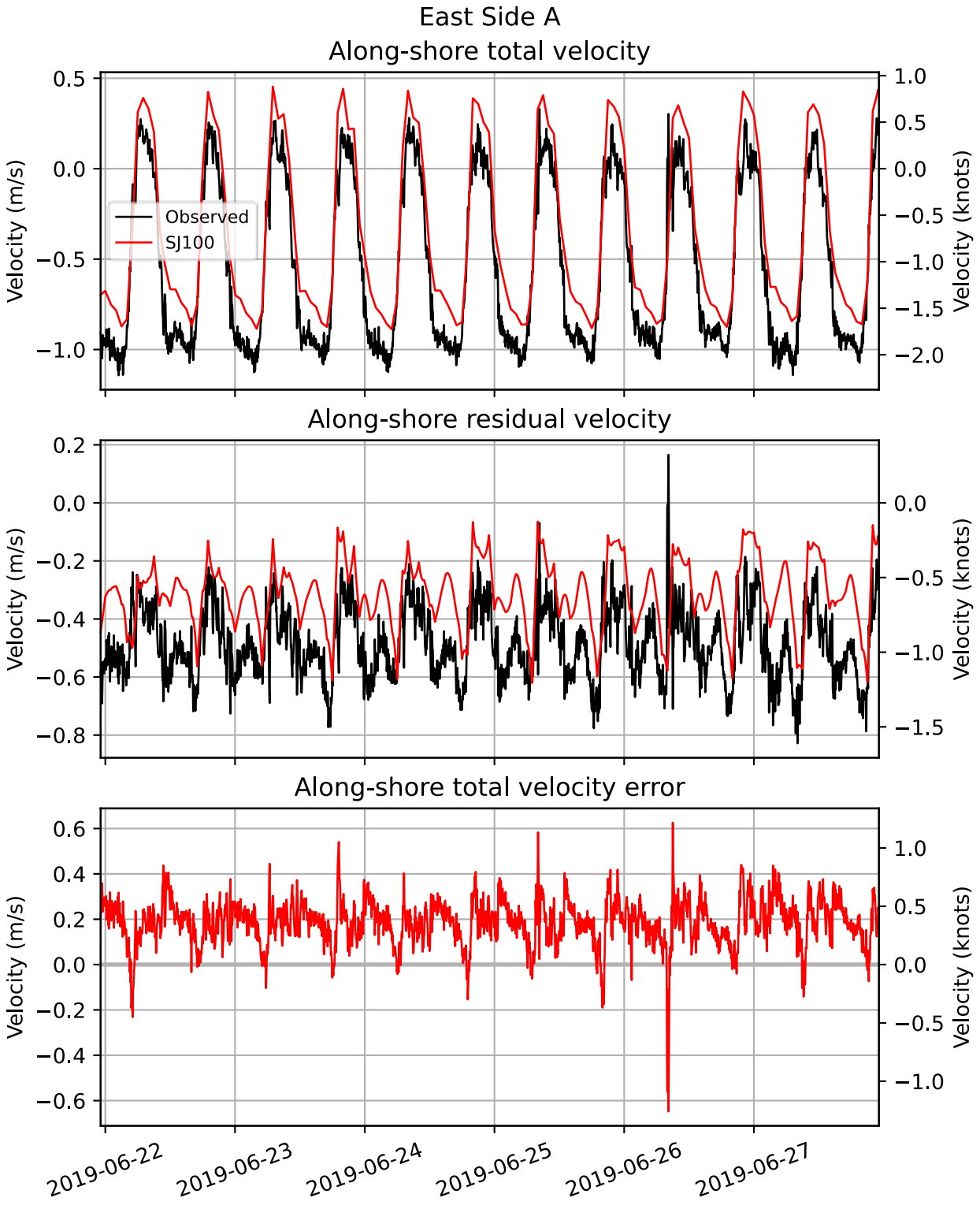

Figure 26: One-week timeseries of along-shore velocity (total and non-tidal) for the eastern HADCP for mid-2019 deployment.

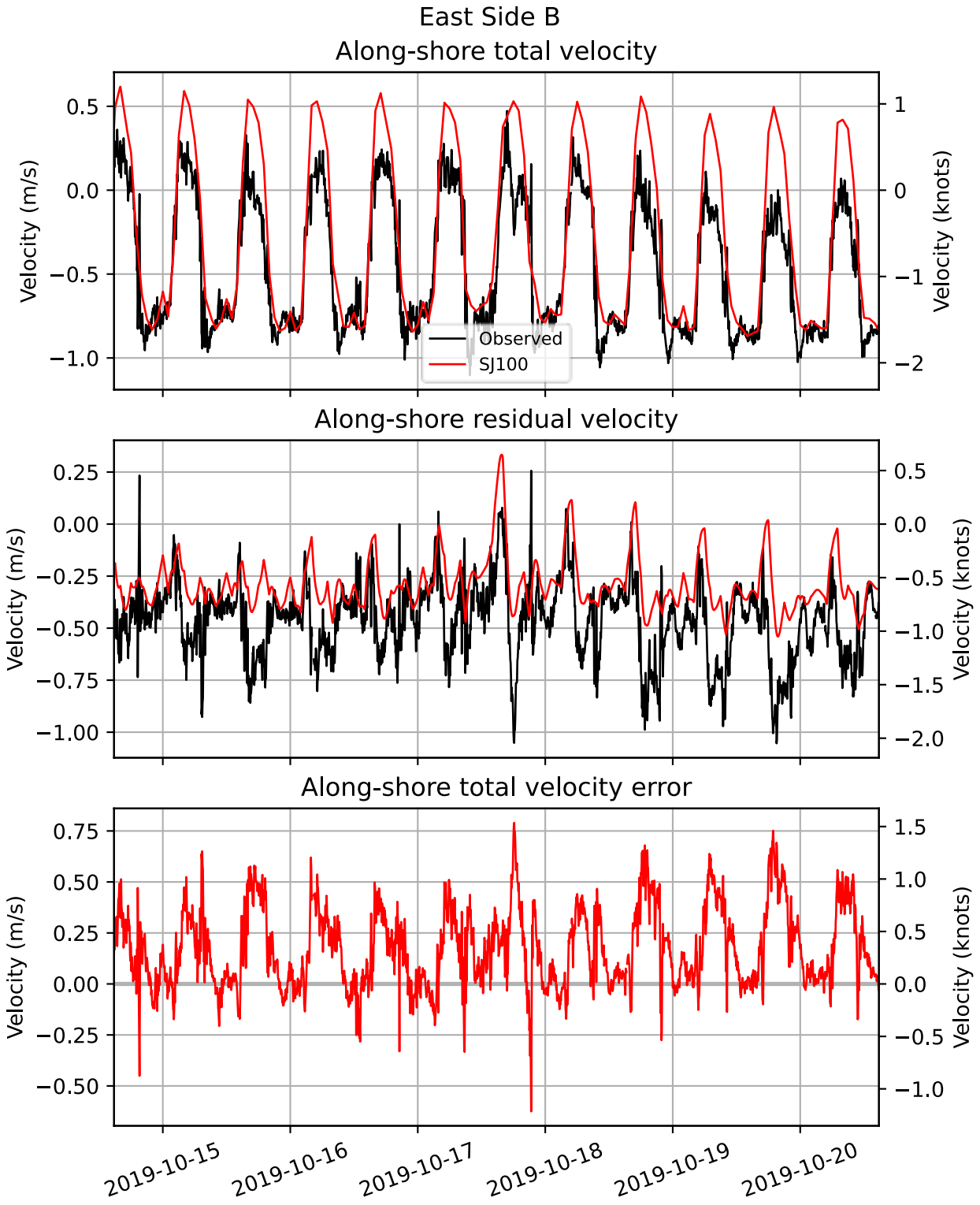

Figure 27: One week time series of along-shore velocity (total and non-tidal) for the eastern HADCP for the fall 2019 deployment.

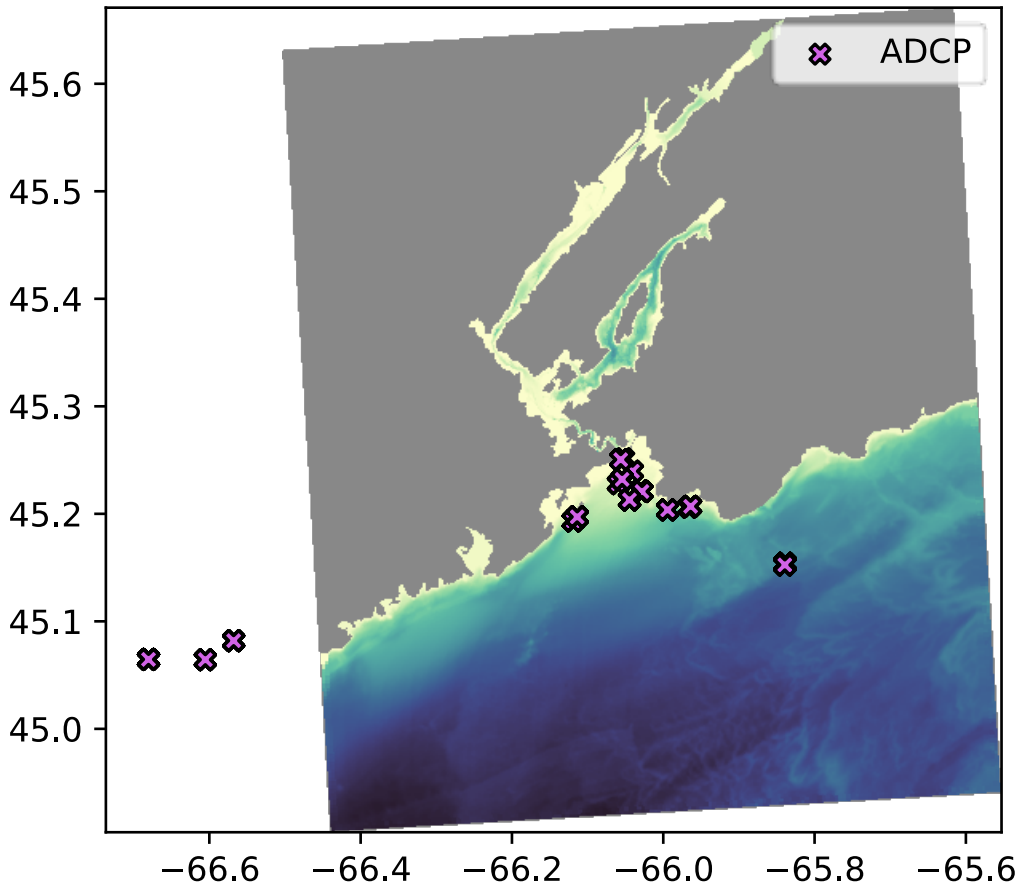

Figure 28: Locations of ADCP observations. The three observations outside the SJ100 domain are in Mace's Bay and are present in Fundy500.

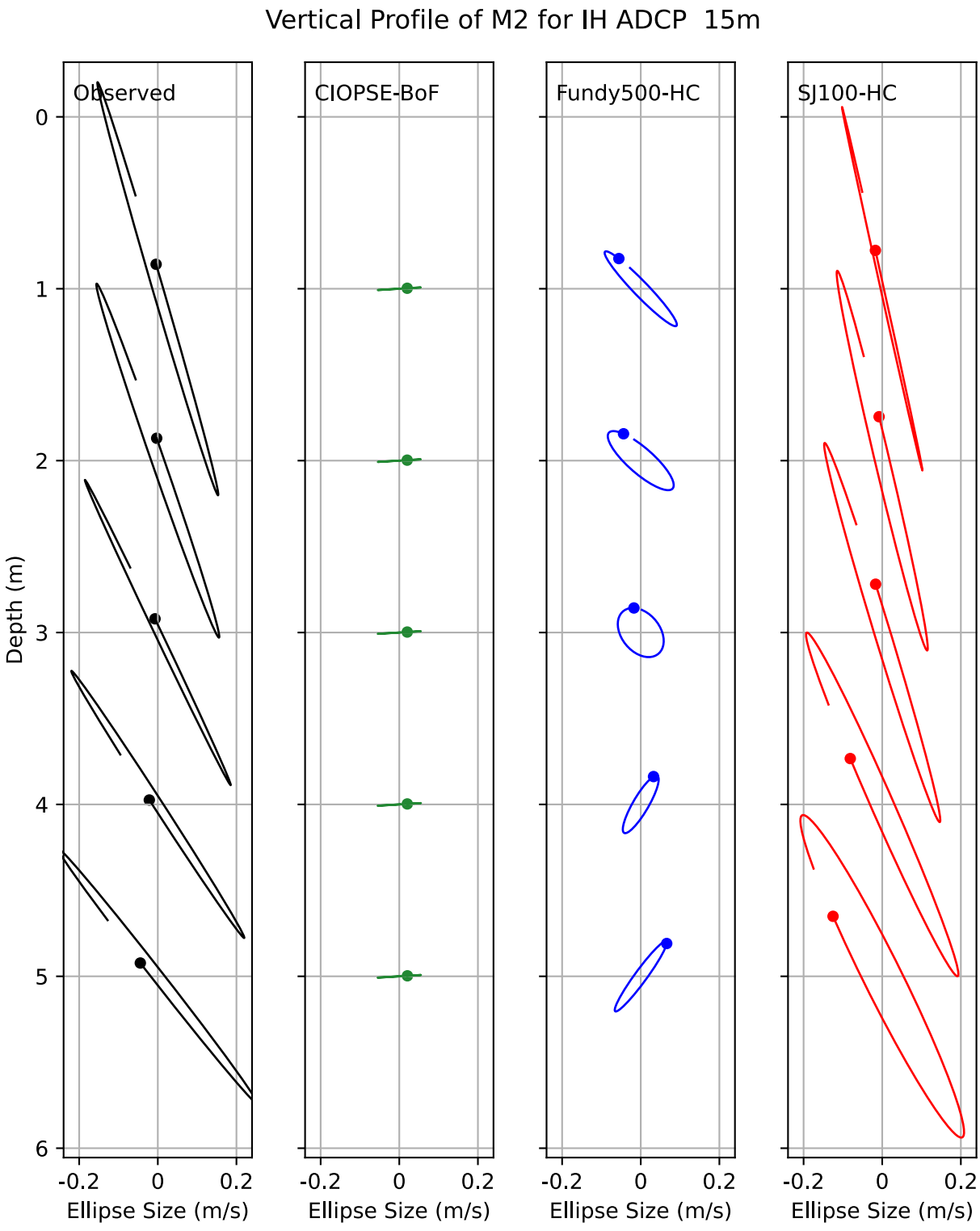

Figure 29: M2 tidal ellipses for Inner Harbour (IH) ADCP.

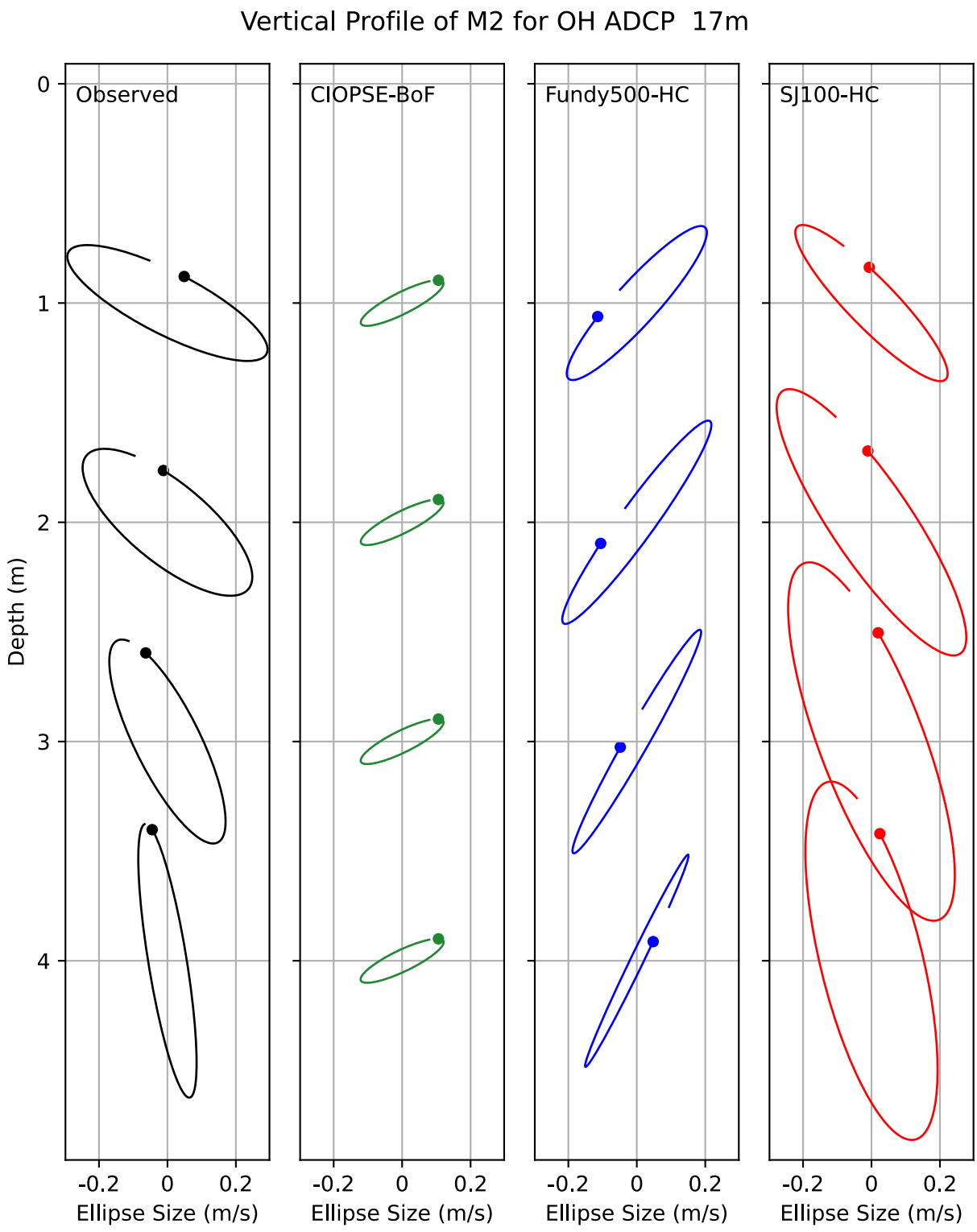

Figure 30: M2 tidal ellipses of Outer Harbour (OH) ADCP.

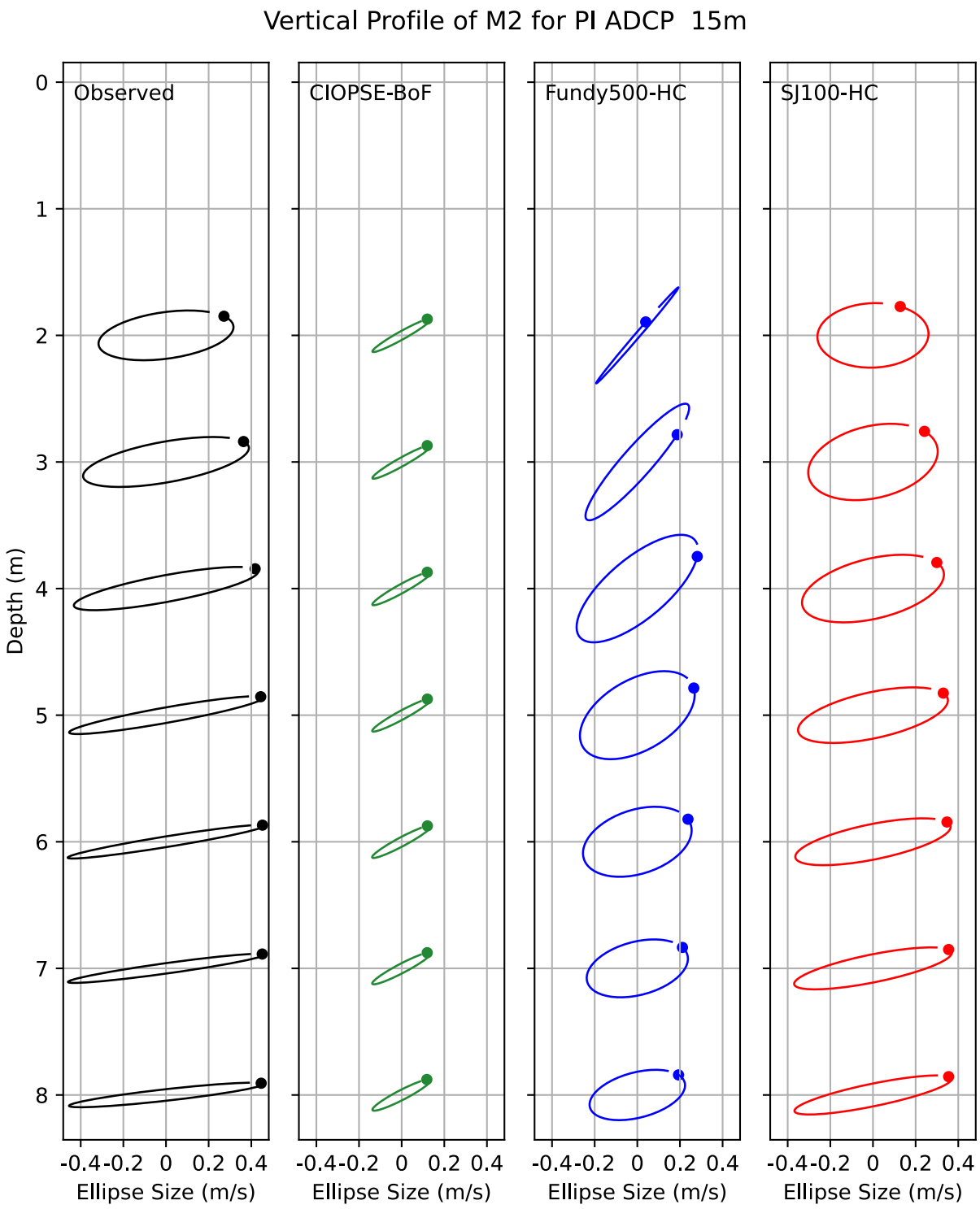

Figure 31: M2 tidal ellipses for Partridge Island (PI) ADCP.

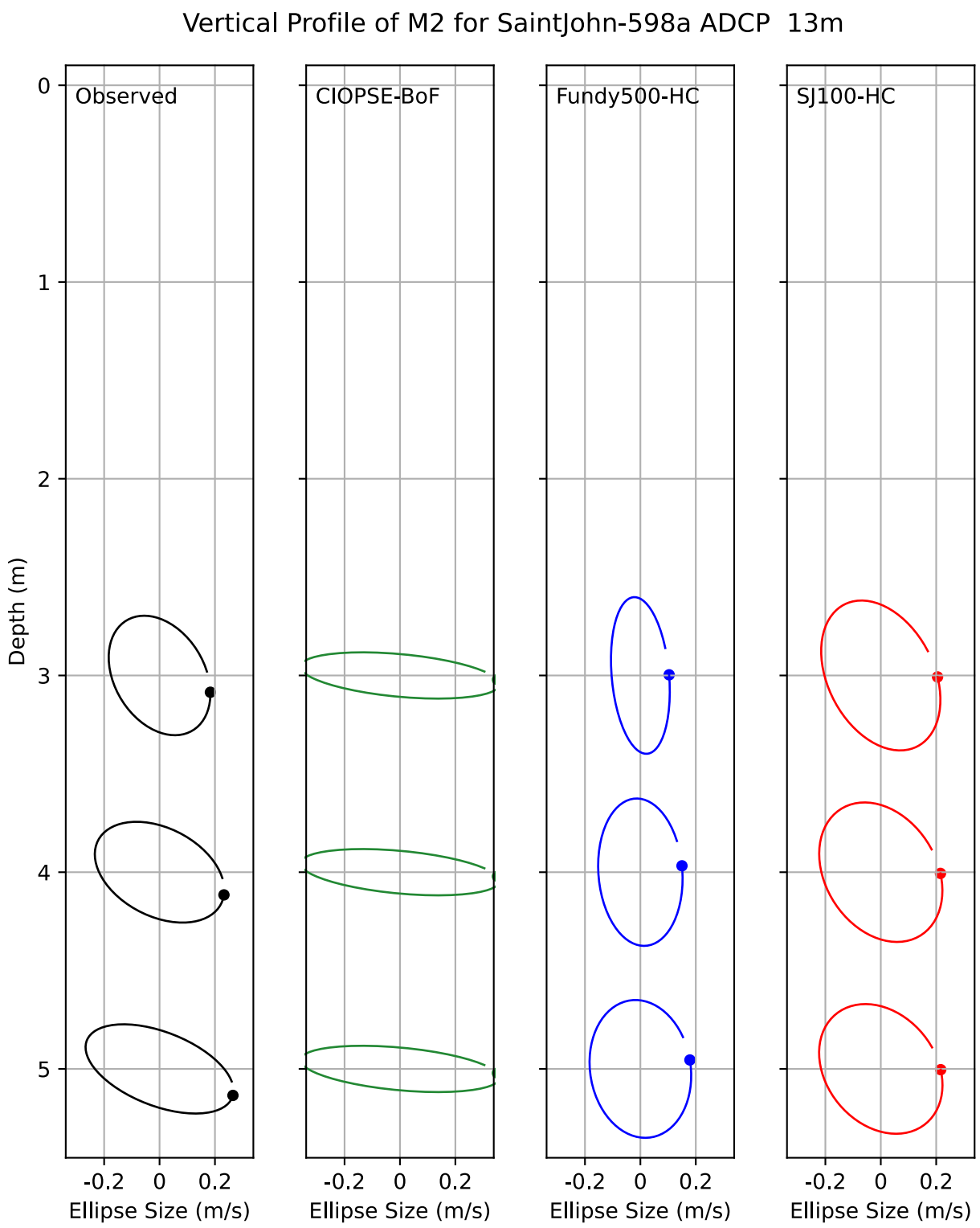

Figure 32: M2 tidal ellipses for ADCP 598, at Canaport.

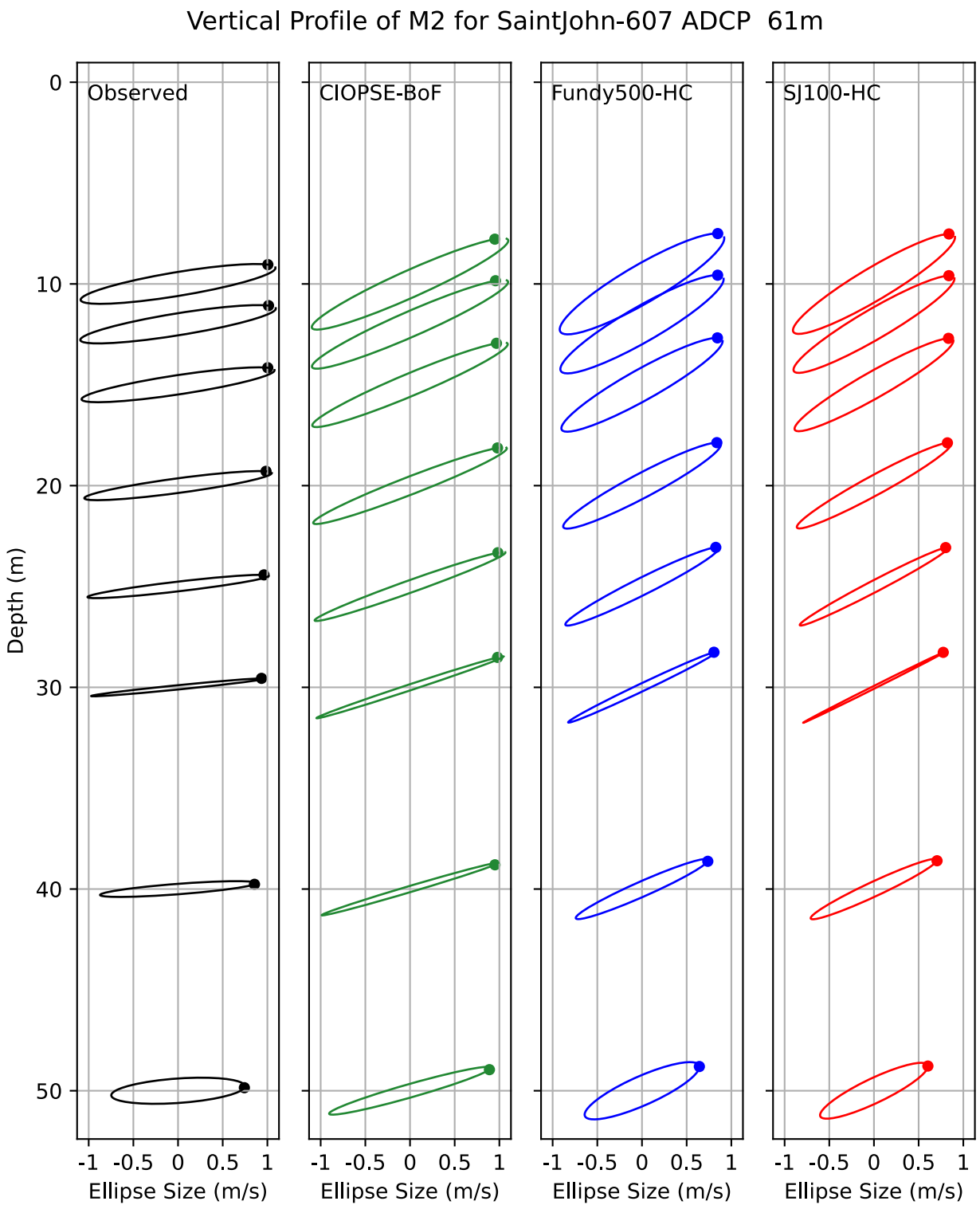

Figure 33: M2 tidal ellipses for ADCP 607, off Cape Spencer.

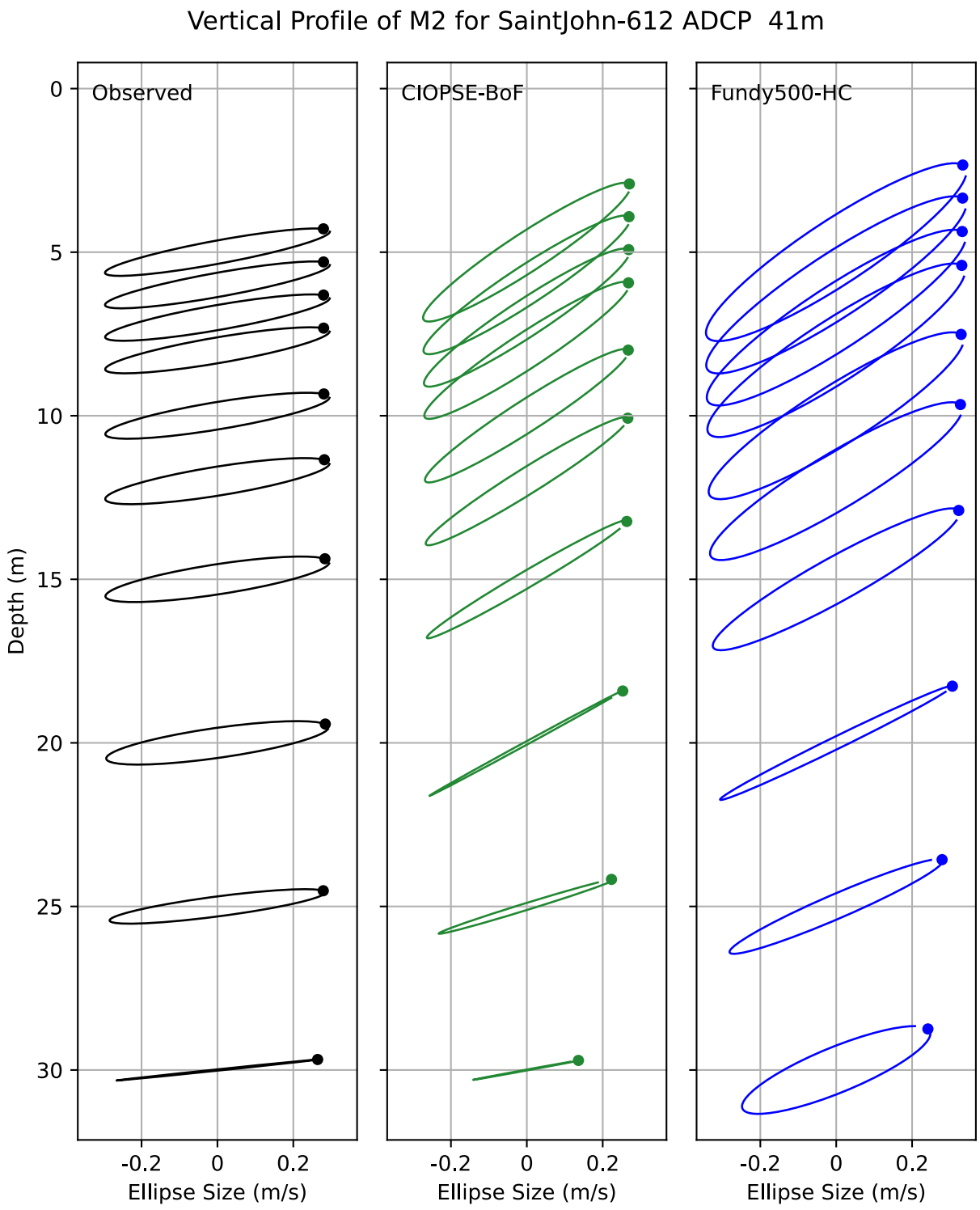

Figure 34: M2 tidal ellipses for ADCP 612, in Mace's Bay (outside of SJ100 domain).

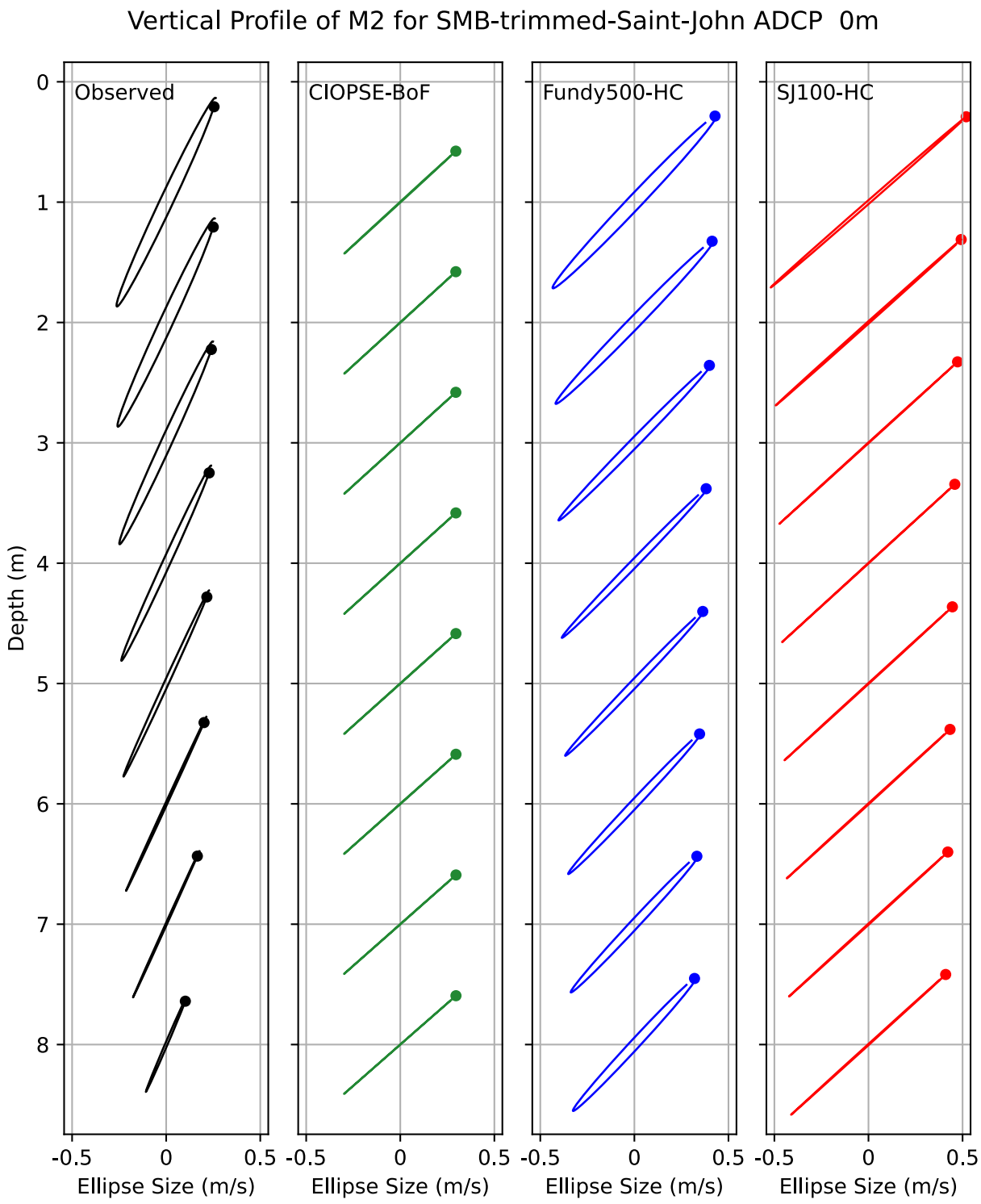

Figure 35: M2 tidal ellipses at the Smart Buoy. This ADCP looks downward from the (heaving) sea surface.

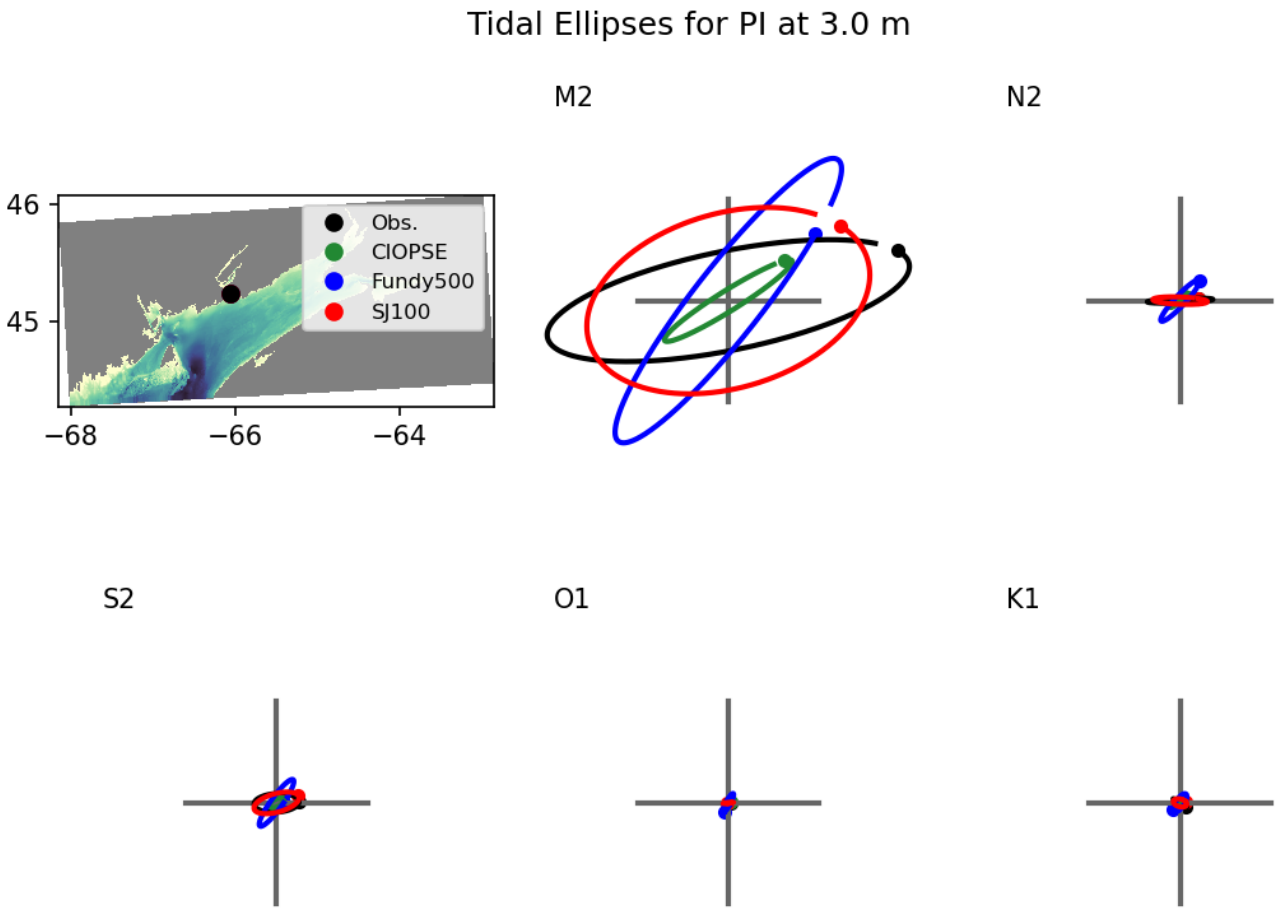

Figure 36: All tidal ellipses at Partridge Island ADCP at 3.0 meters below the surface. Relative scale of S2, N2, and diurnal constituents to M2 ellipses is similar for all other ADCPs in the domain.

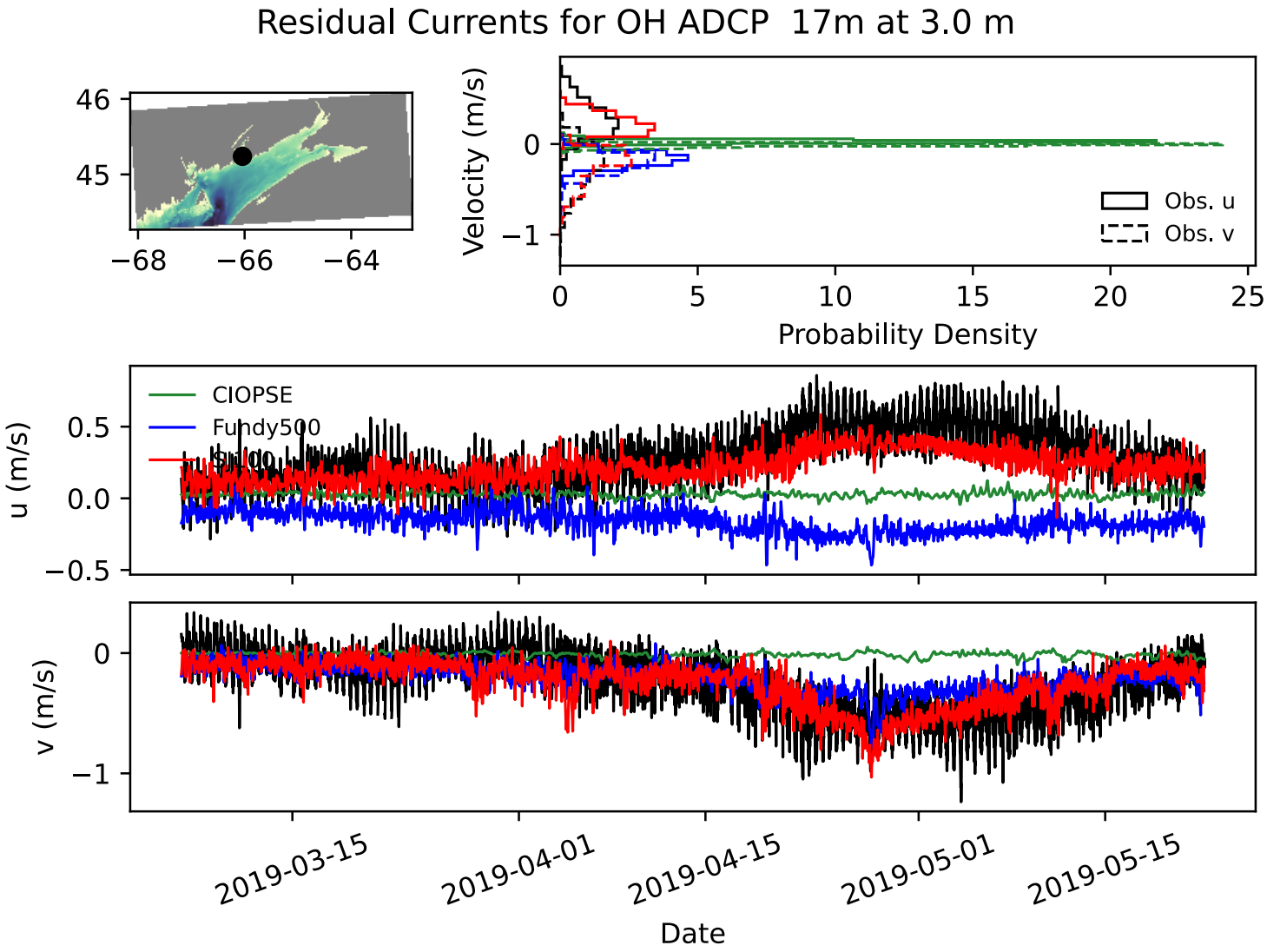

Figure 37: Non-tidal currents at Outer Harbour at 3.0 meters below the surface.

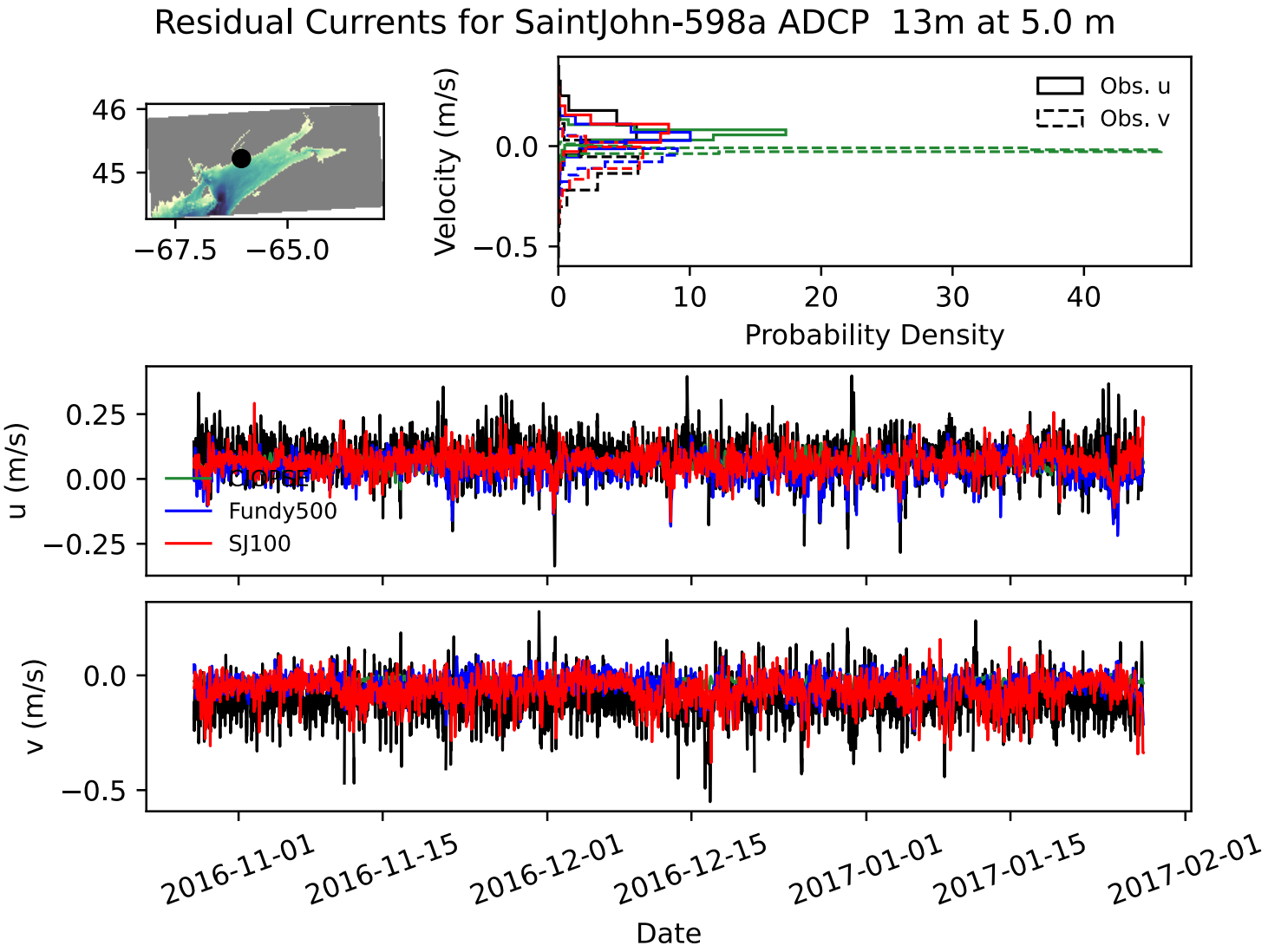

Figure 38: Non-tidal currents at ADCP 598 at 5.0 meters below surface.

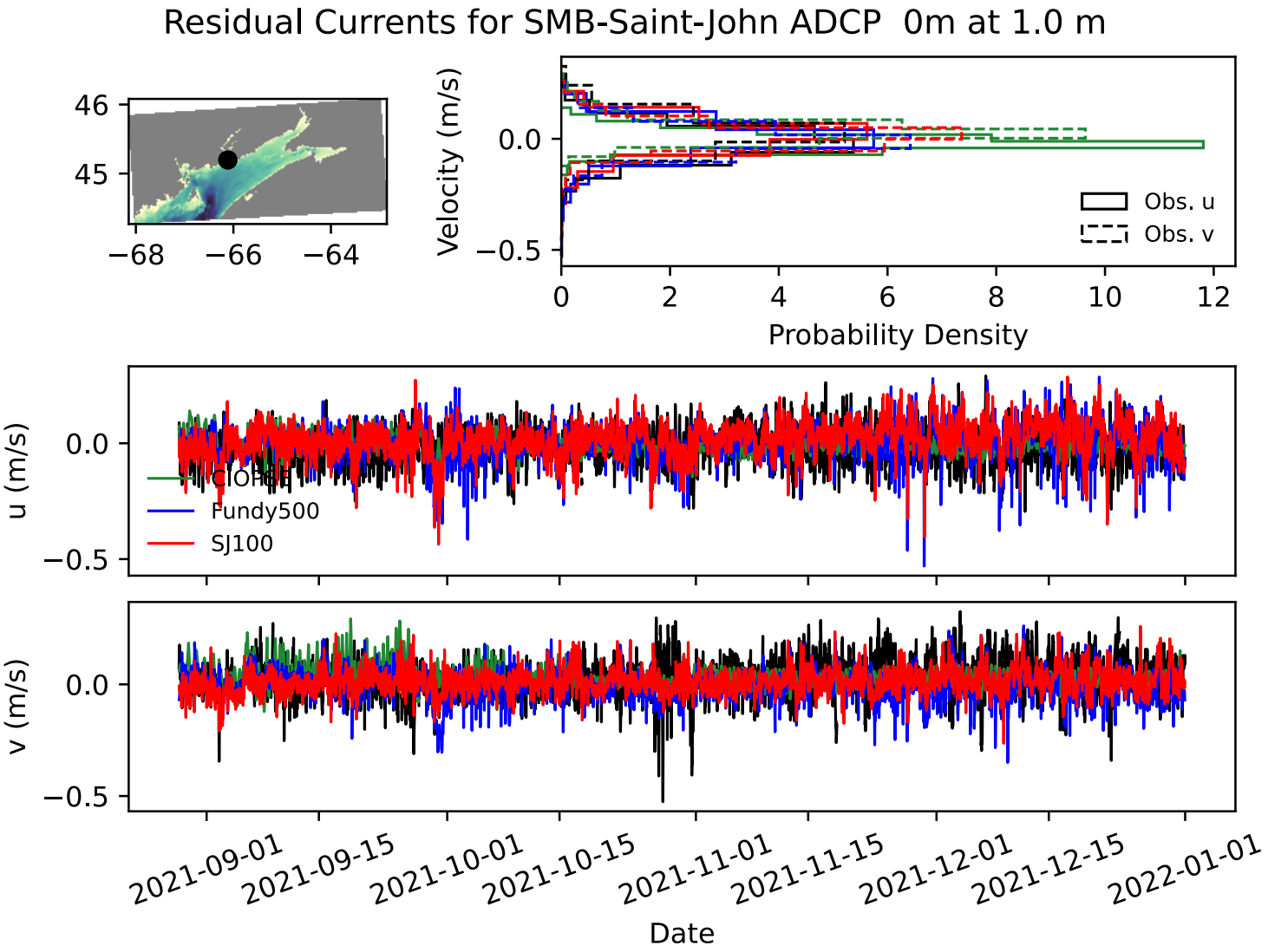

Figure 39: Near-surface non-tidal currents at downward-looking ADCP on the Smart Buoy.

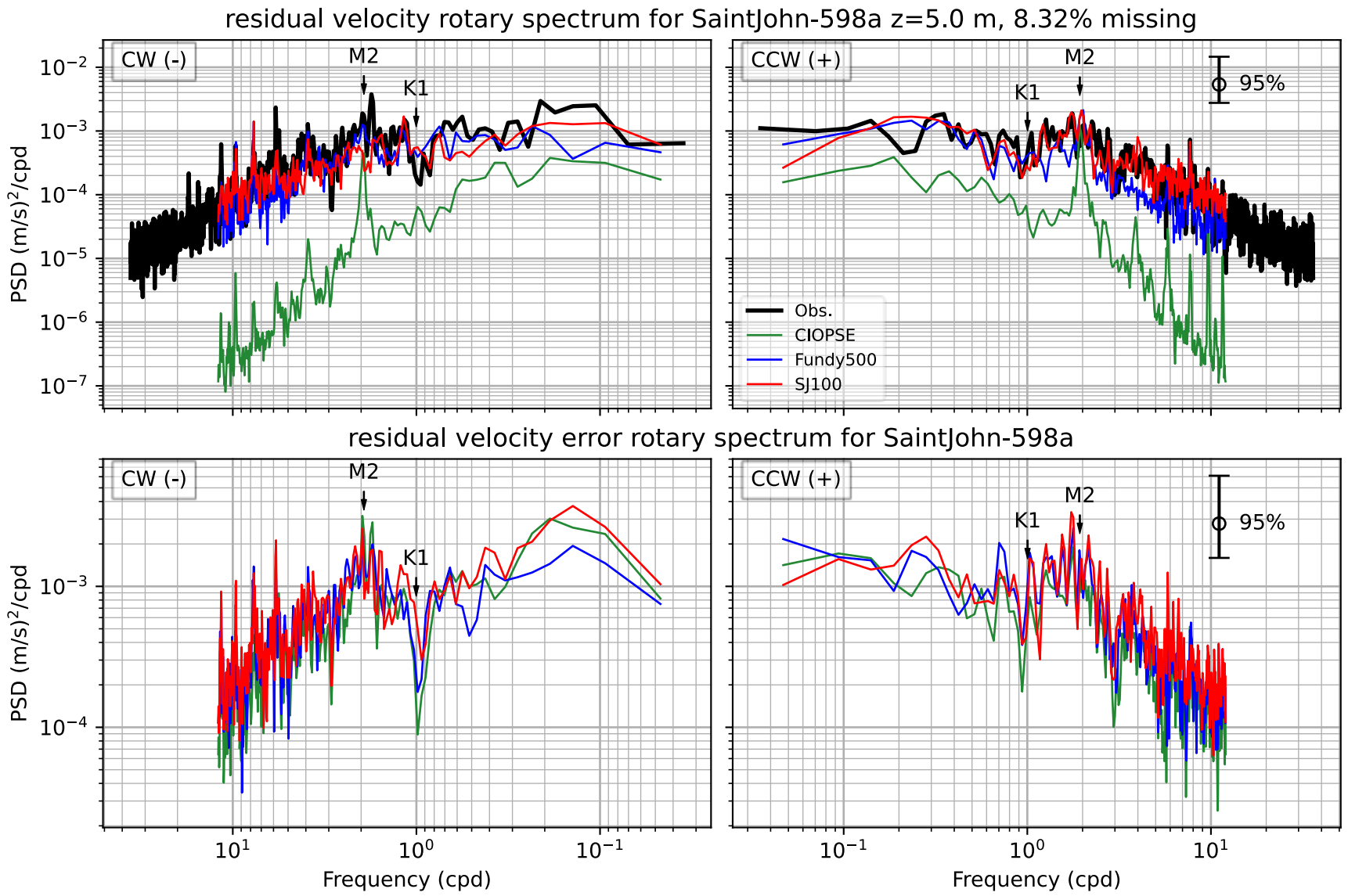

Figure 40: Non-tidal rotary spectra for ADCP 598 at 5.0 m below the surface.

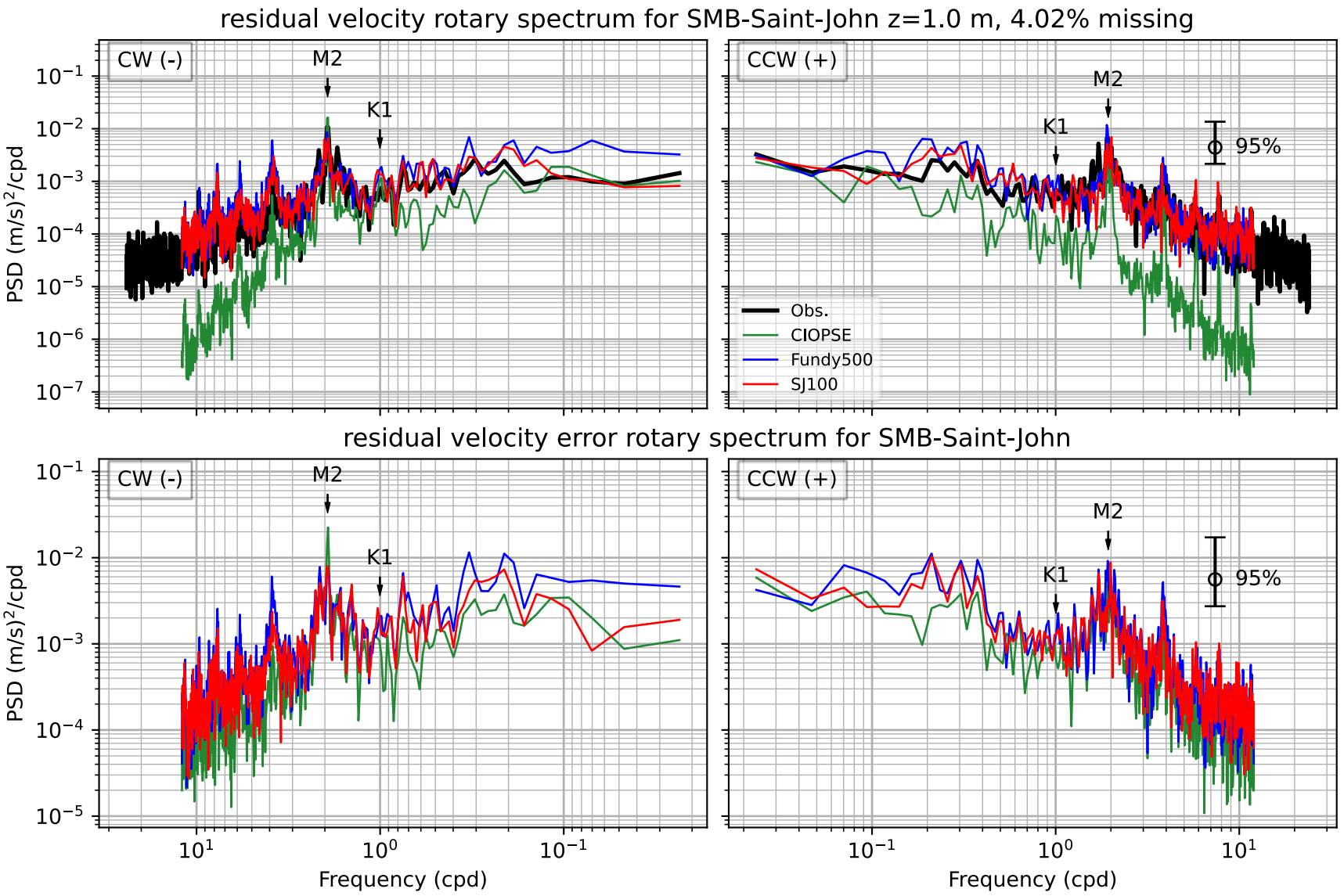

Figure 41: Non-tidal rotary spectra for the downward looking ADCP on the Smart Buoy at 1.0 m below the surface.

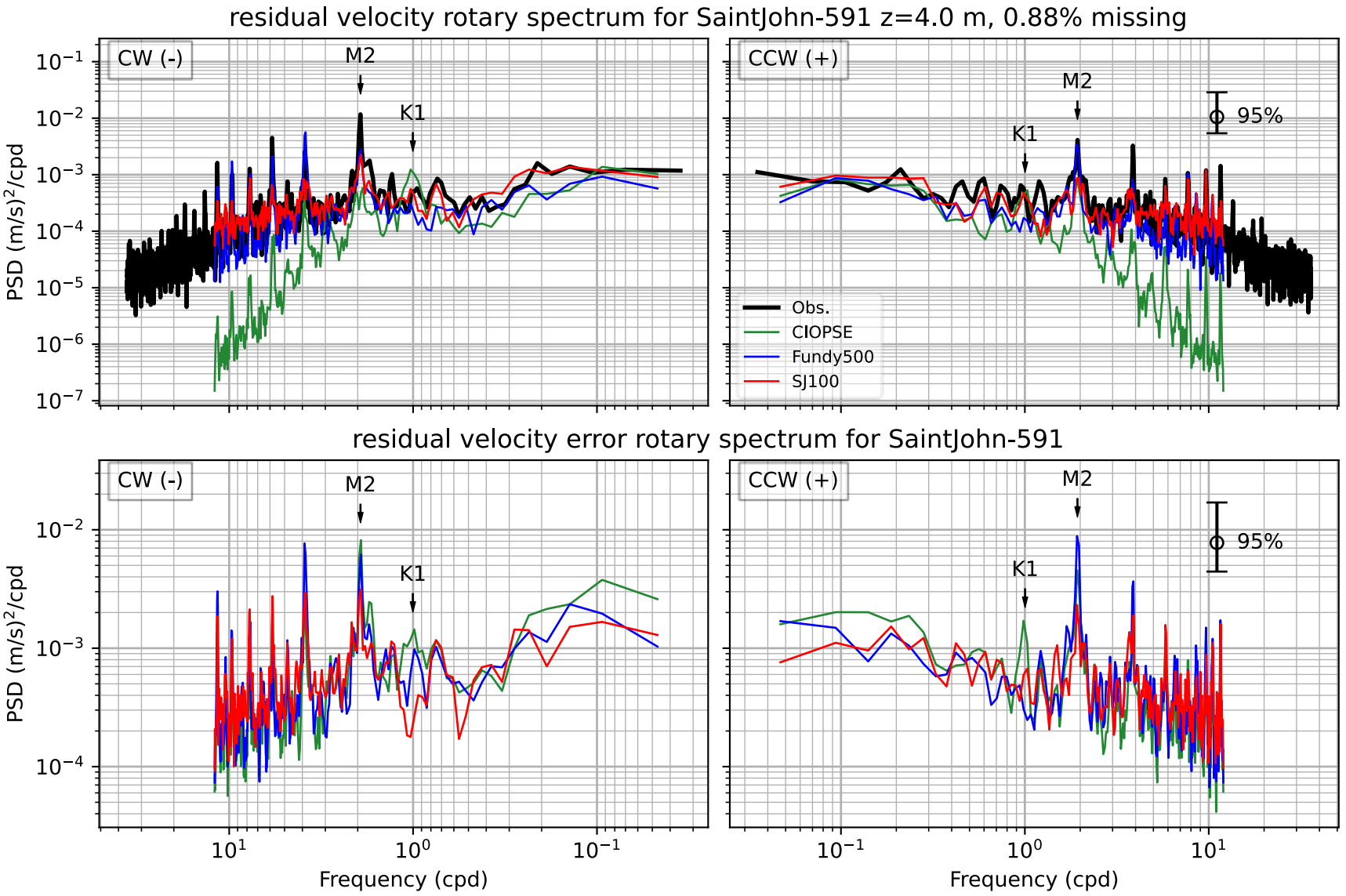

Figure 42: Non-tidal rotary spectra for ADCP 591 at 4.0 m below the surface.

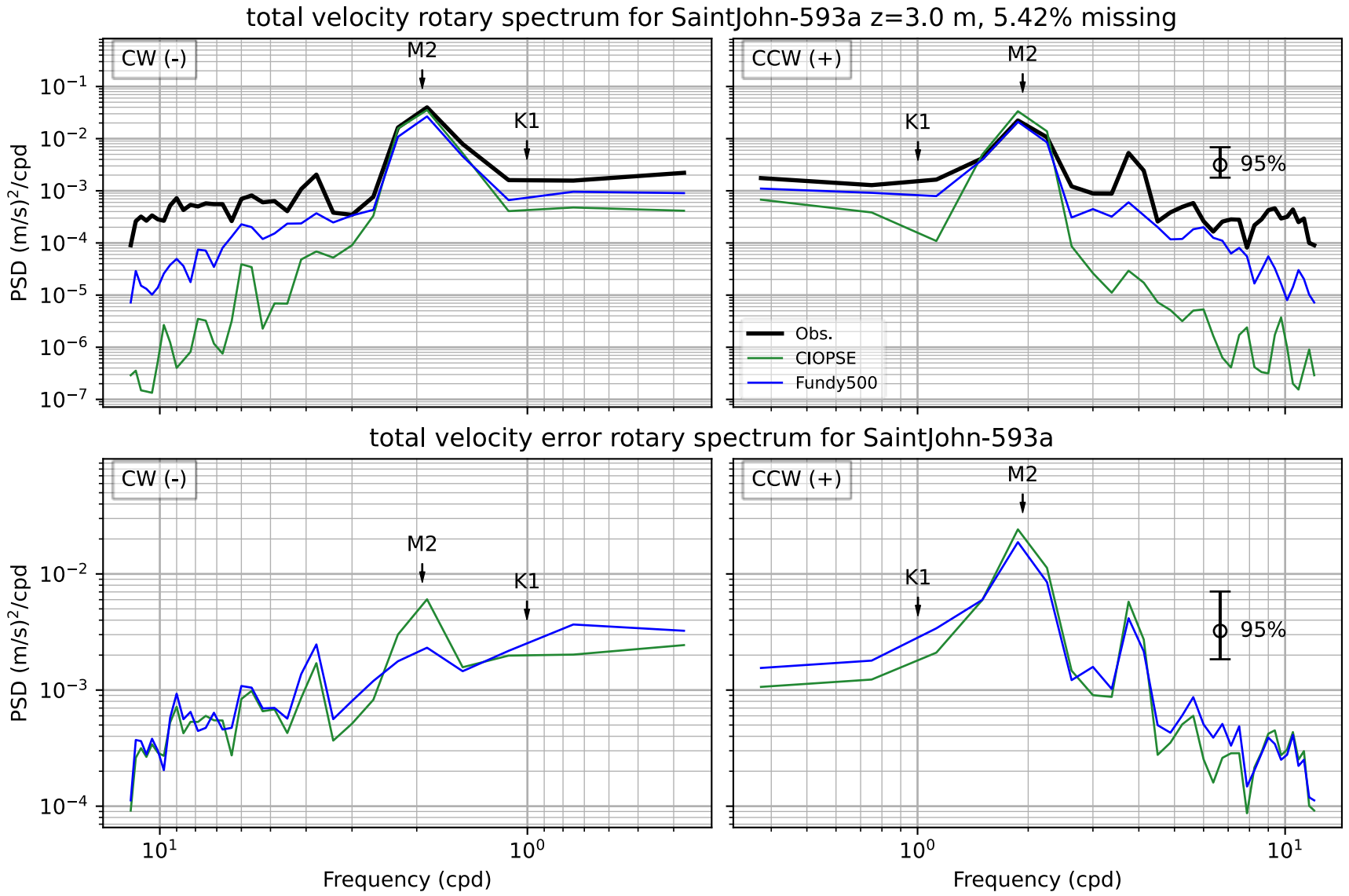

Figure 43: Non-tidal rotary spectra for ADCP 593 (in Mace's Bay) at 3.0 m below the surface.

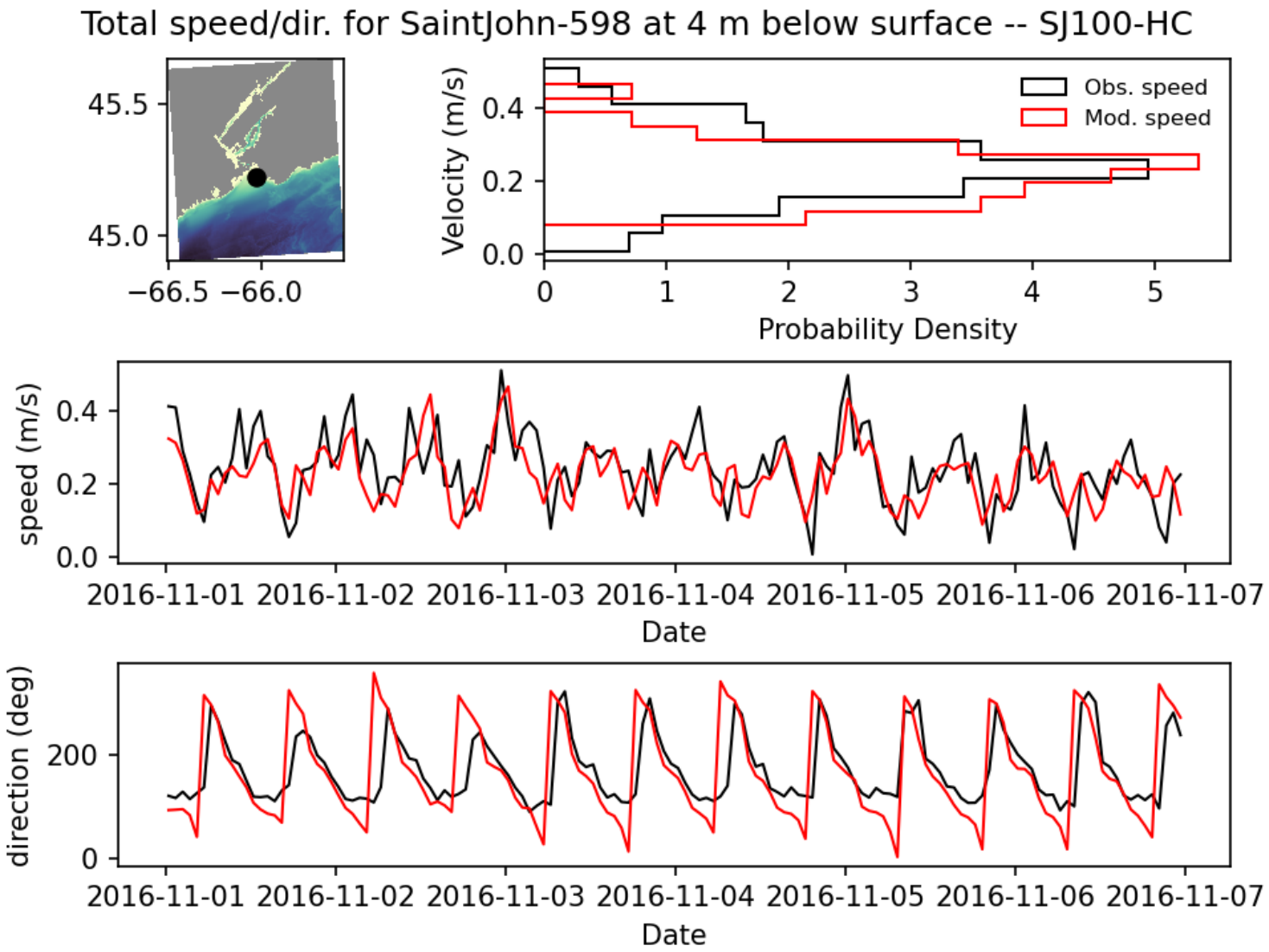

Figure 44: One week of total velocity for ADCP 598 (at Canaport).

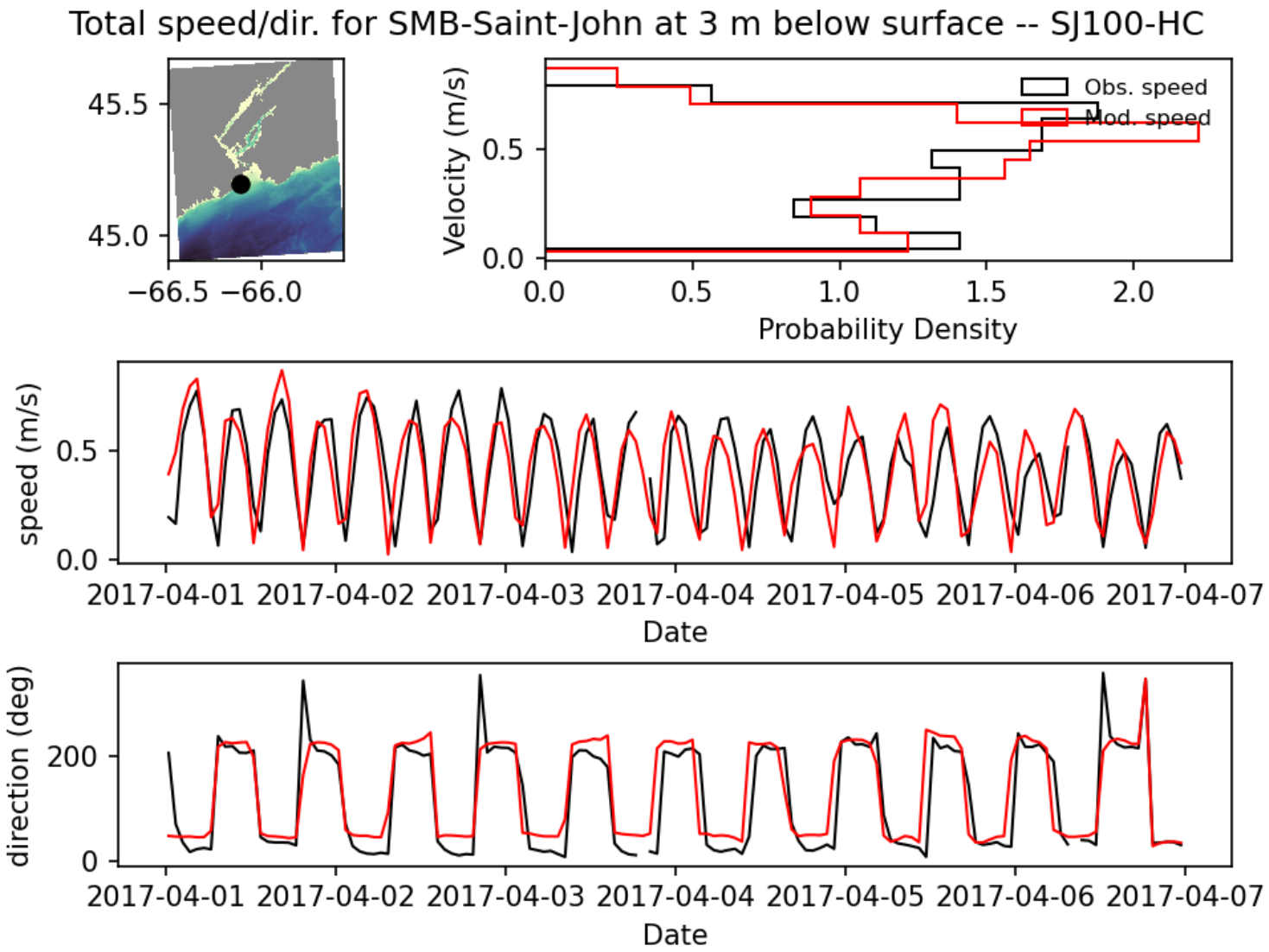

Figure 45: One week of total velocity for the Smart Buoy ADCP.

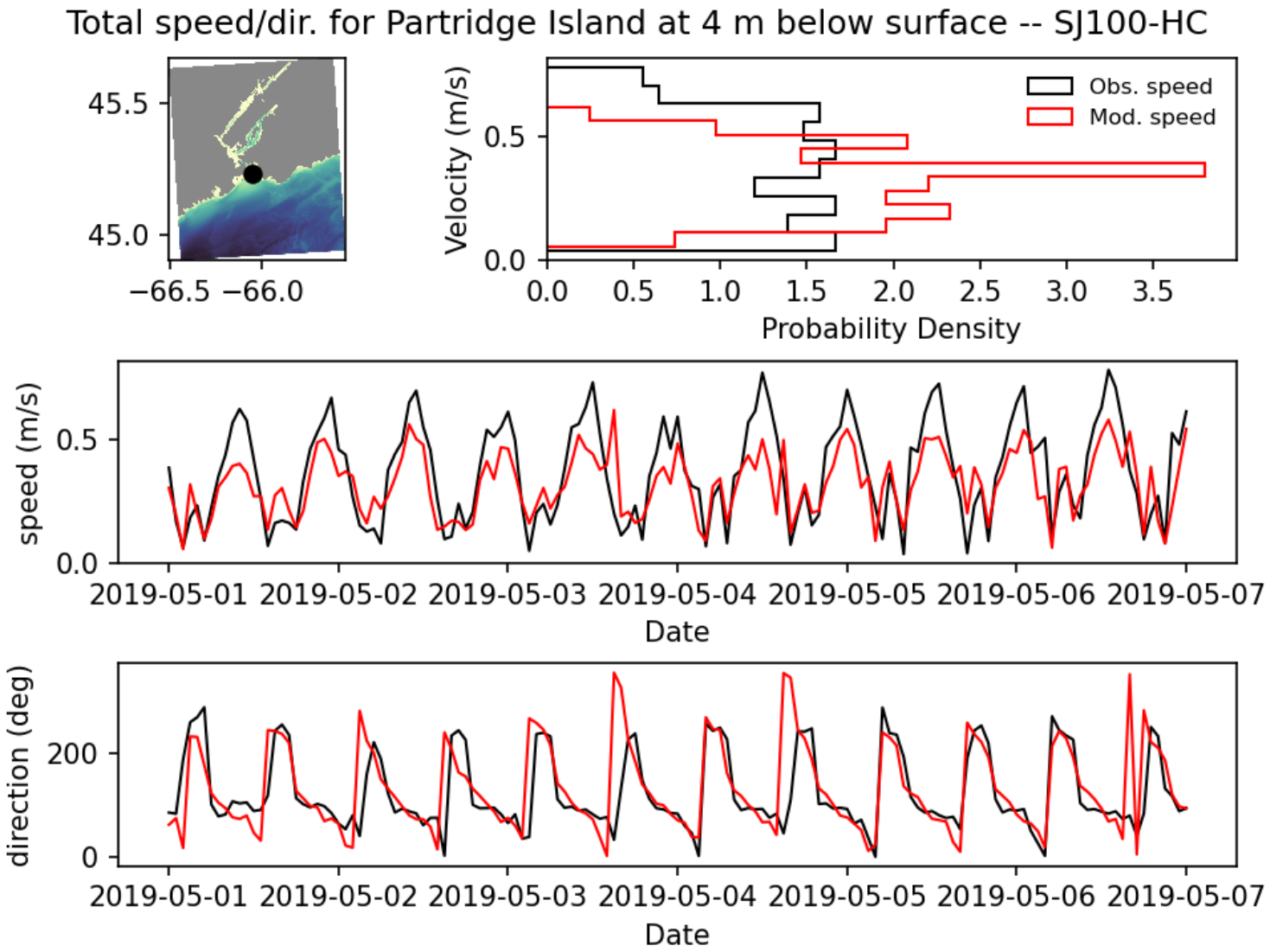

Figure 46: One week of total velocity for Partridge Island ADCP.

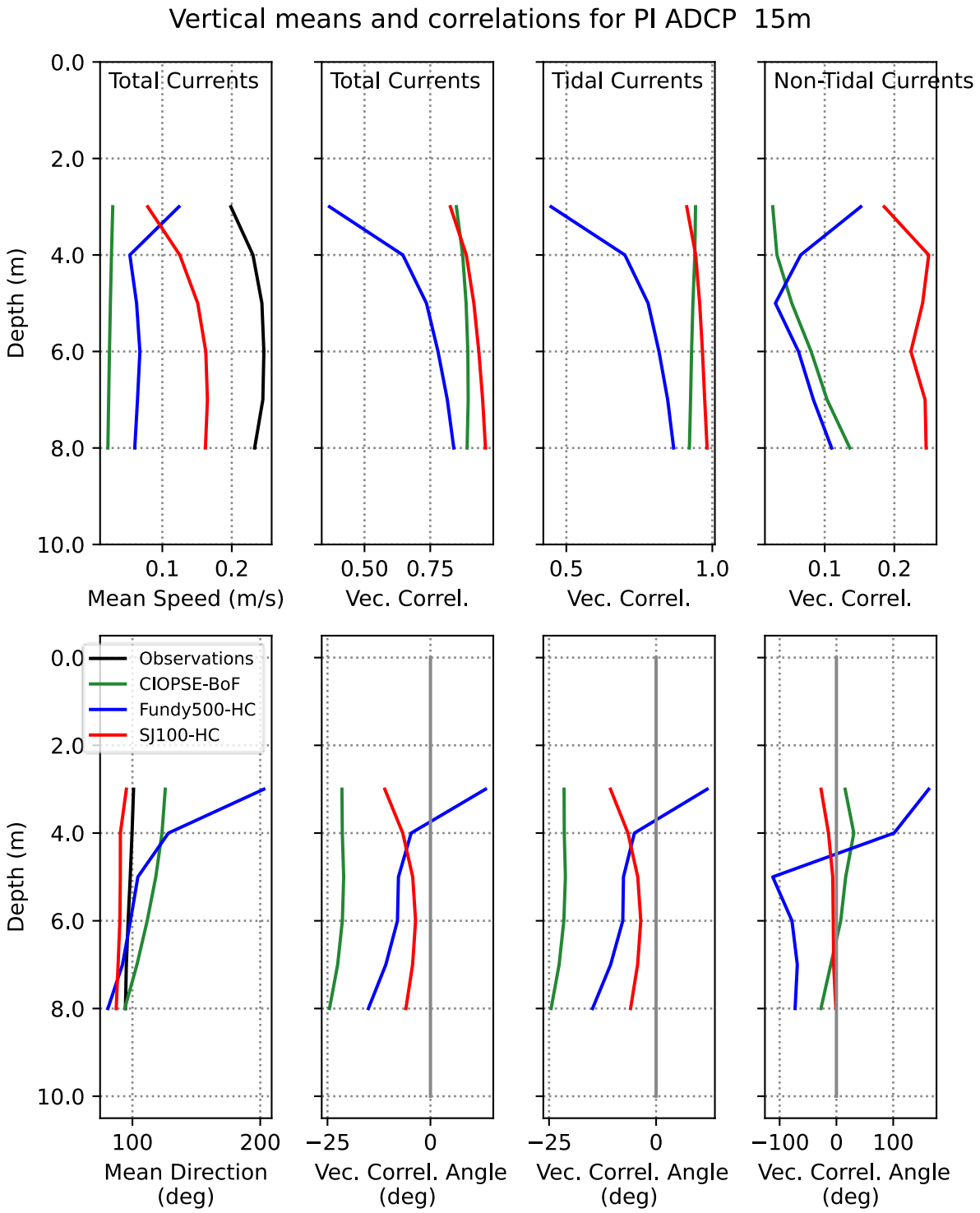

Figure 47: Depth profiles of mean current speeds (upper panels) and directions (lower panels) and related vector correlations for Partridge Island (PI) ADCP.

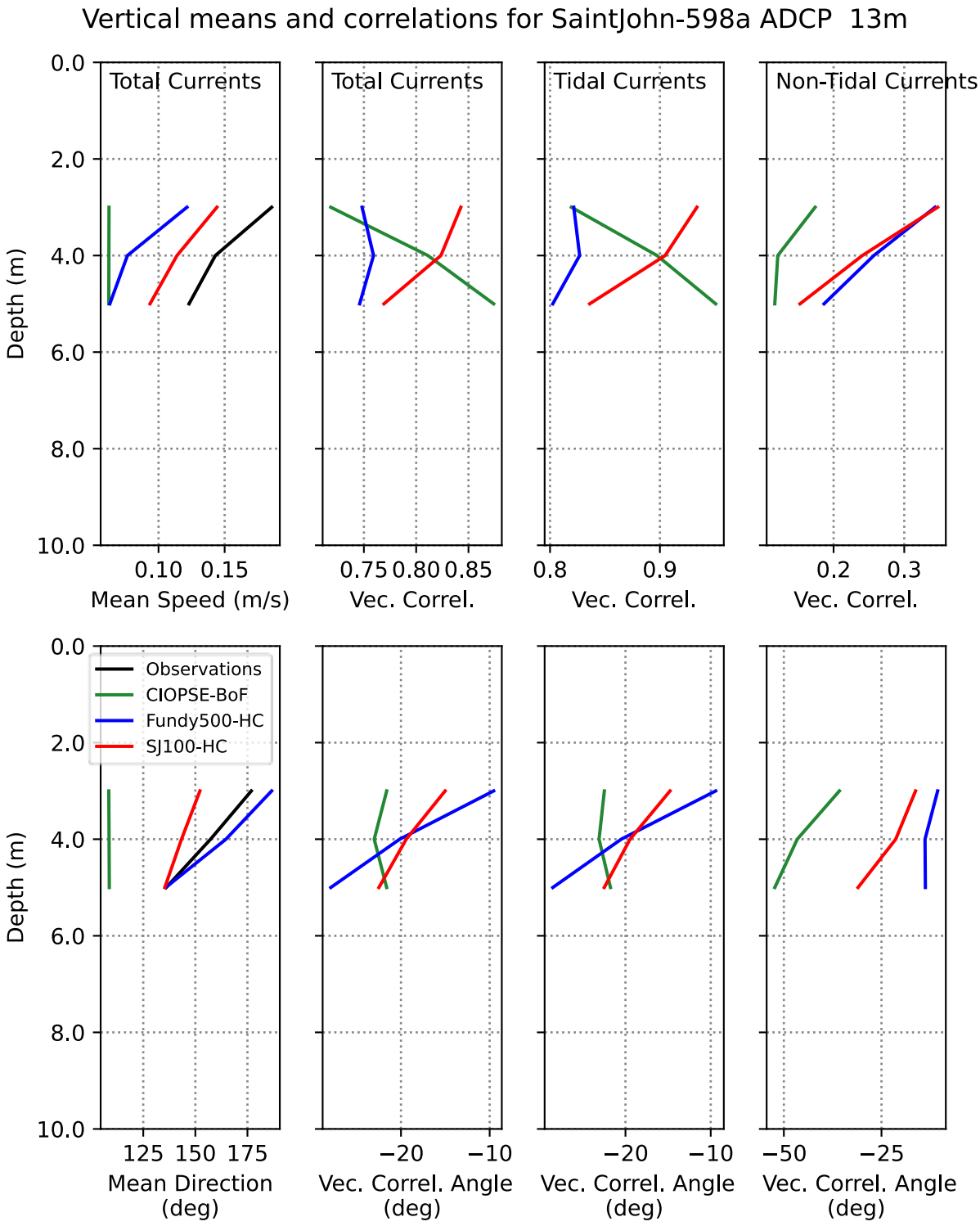

Figure 48: Depth profiles of mean current speeds (upper panels) and directions (lower panels) and related vector correlations for ADCP 598 (at Canaport).

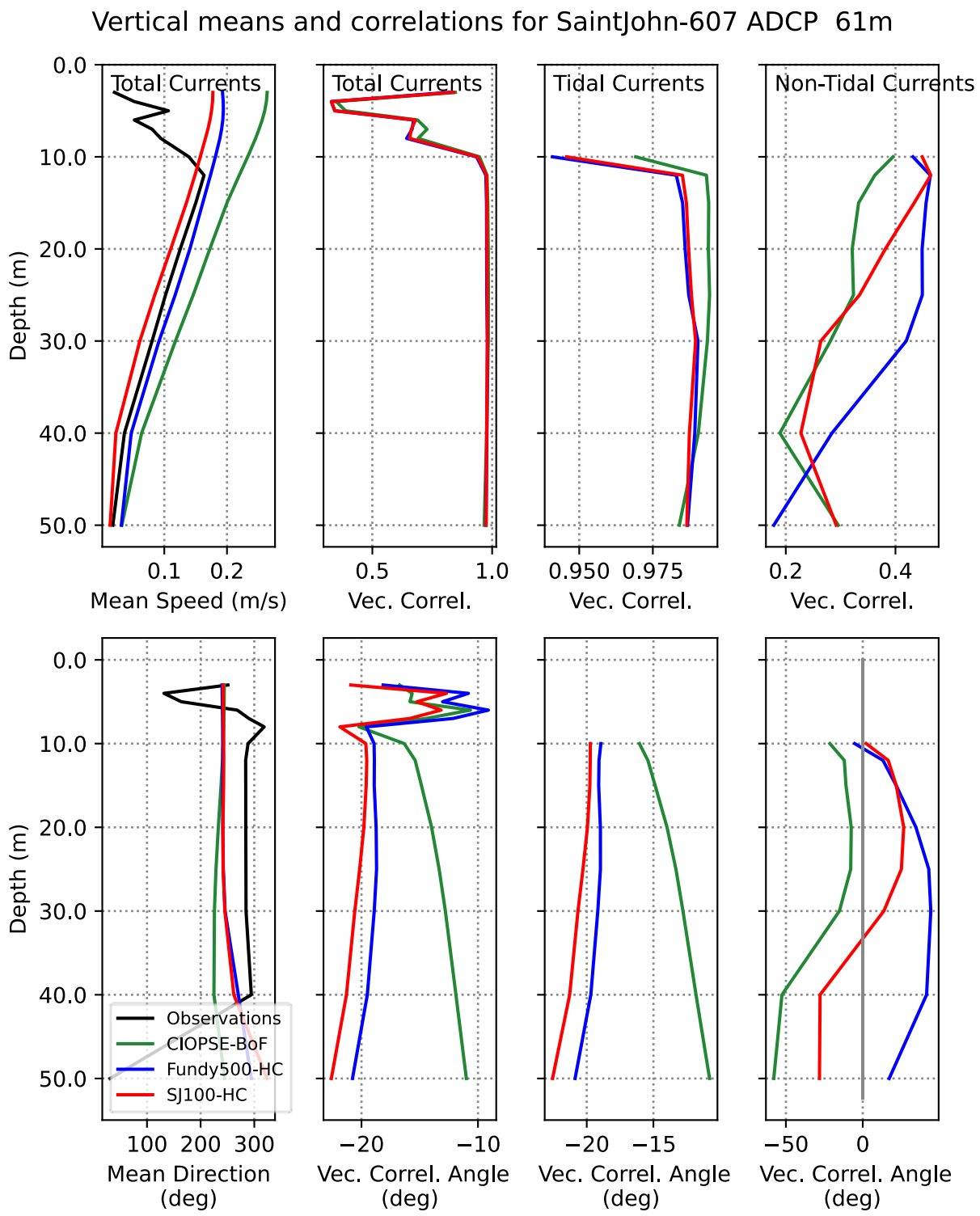

Figure 49: Depth profiles of mean current speeds (upper panels) and directions (lower panels) and related vector correlations for ADCP 607 (off Cape Spencer).

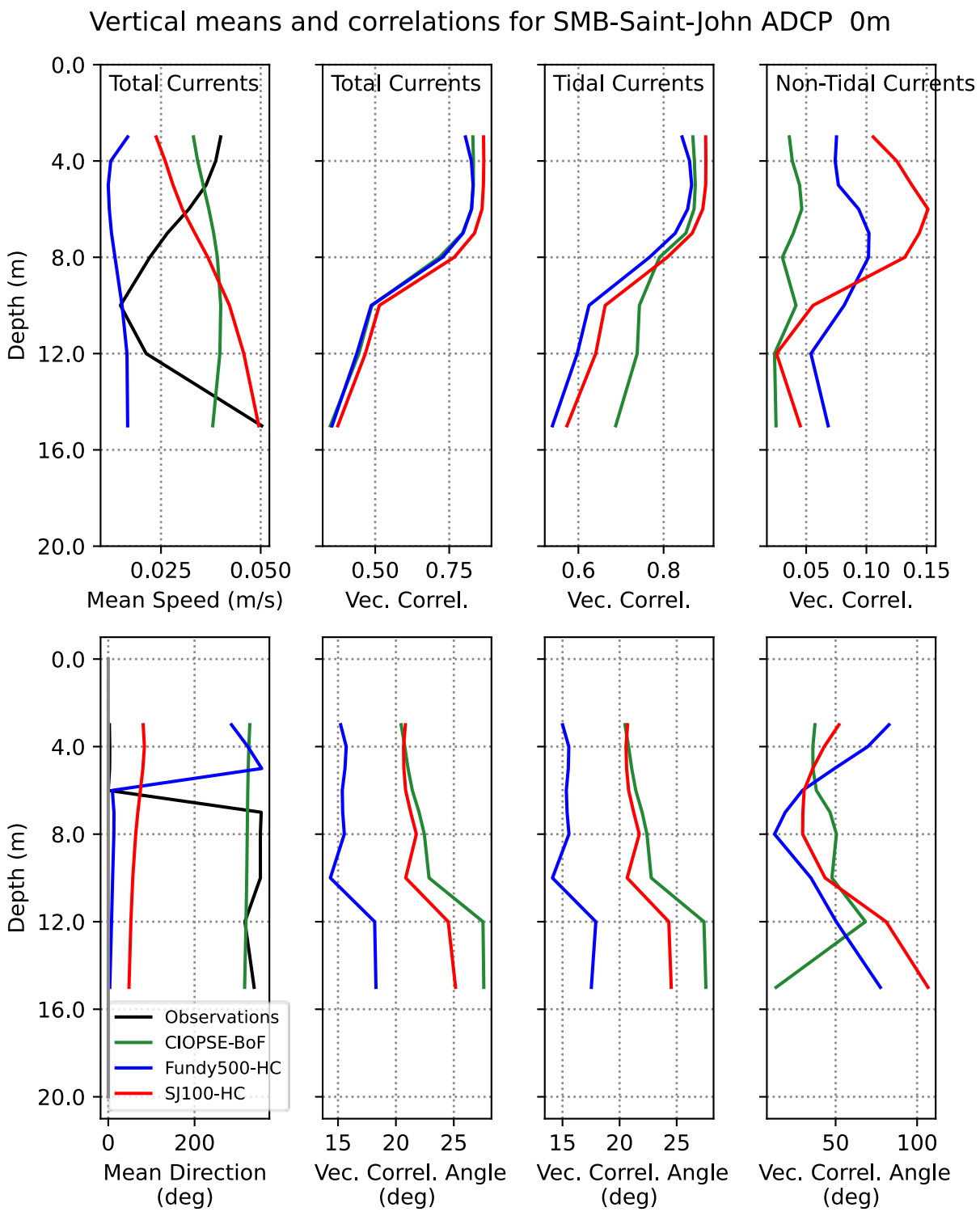

Figure 50: Depth profiles of mean current speeds (upper panels) and directions (lower panels) and related vector correlations for the Smart Buoy ADCP.

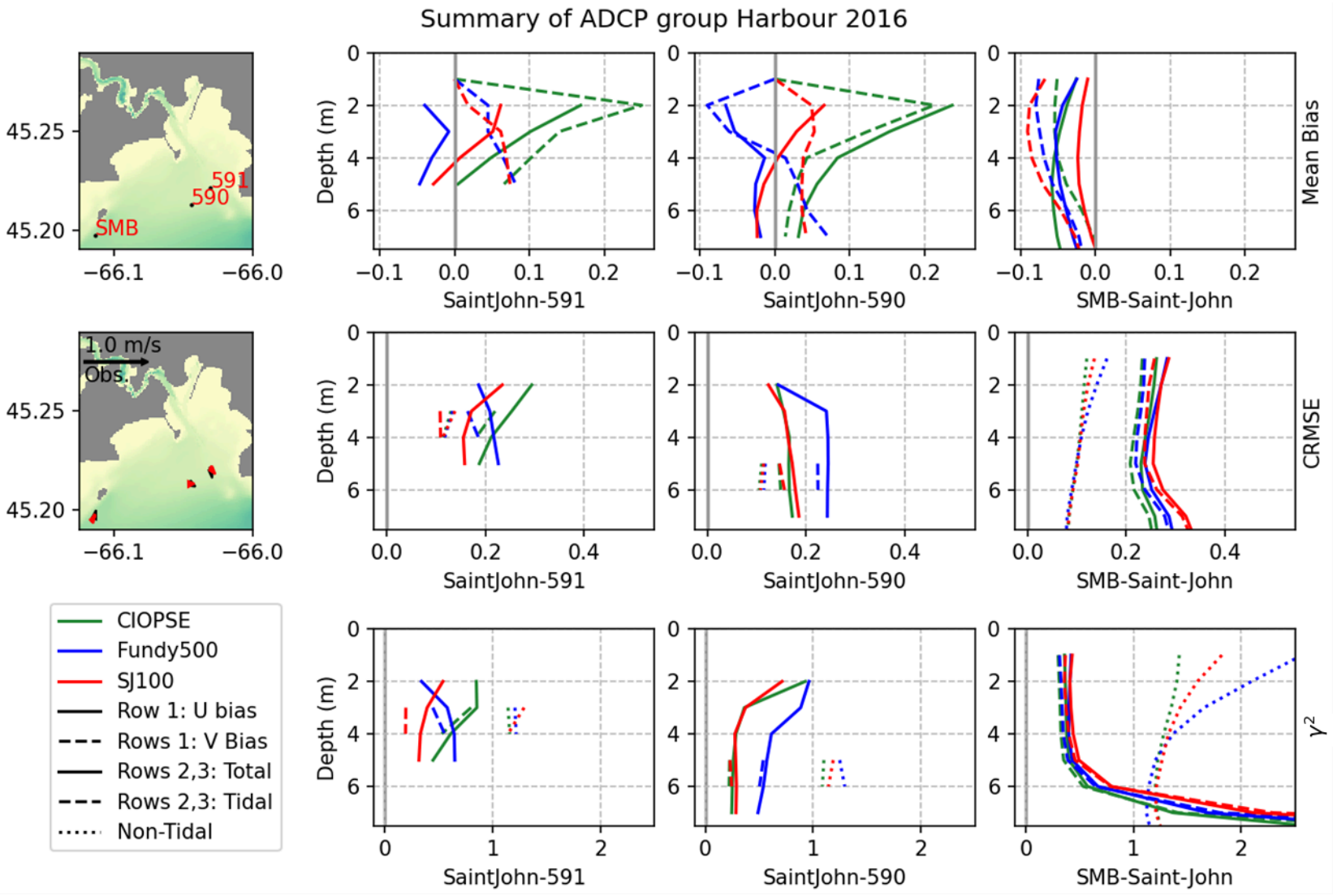

Figure 51: Depth profiles of statistics for moorings near Saint John in 2016. Top row shows mean bias for u (solid) and v (dashed) velocities. Middle and bottom row show CRMSE and γ^2 for total (solid), tidal (dashed), and non-tidal (dotted) circulation.

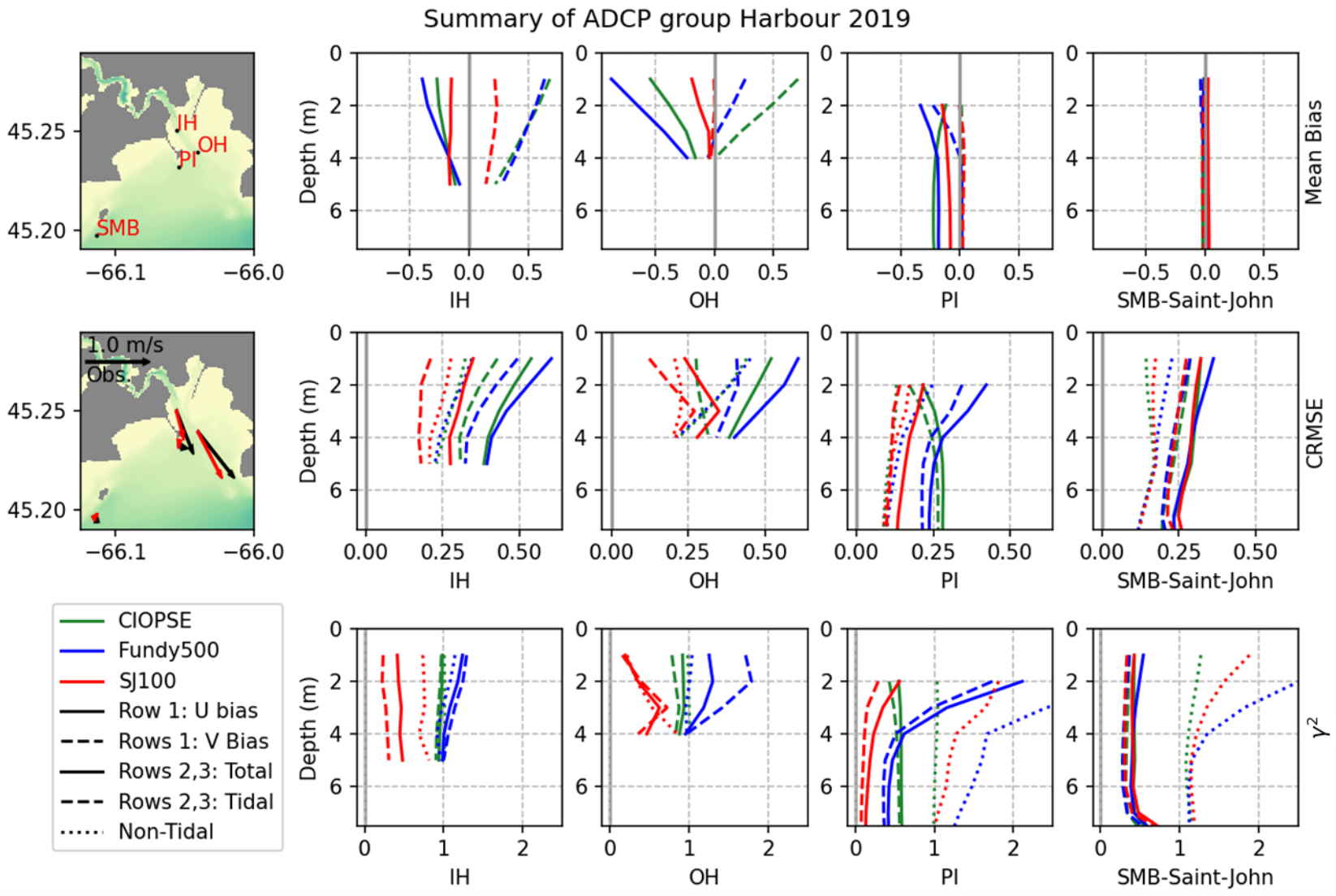

Figure 52: Depth profiles of statistics for moorings in the harbour in 2019. Top row shows mean bias for u (solid) and v (dashed) velocities. Middle and bottom row show CRMSE and γ^2 for total (solid), tidal (dashed), and non-tidal (dotted) circulation.

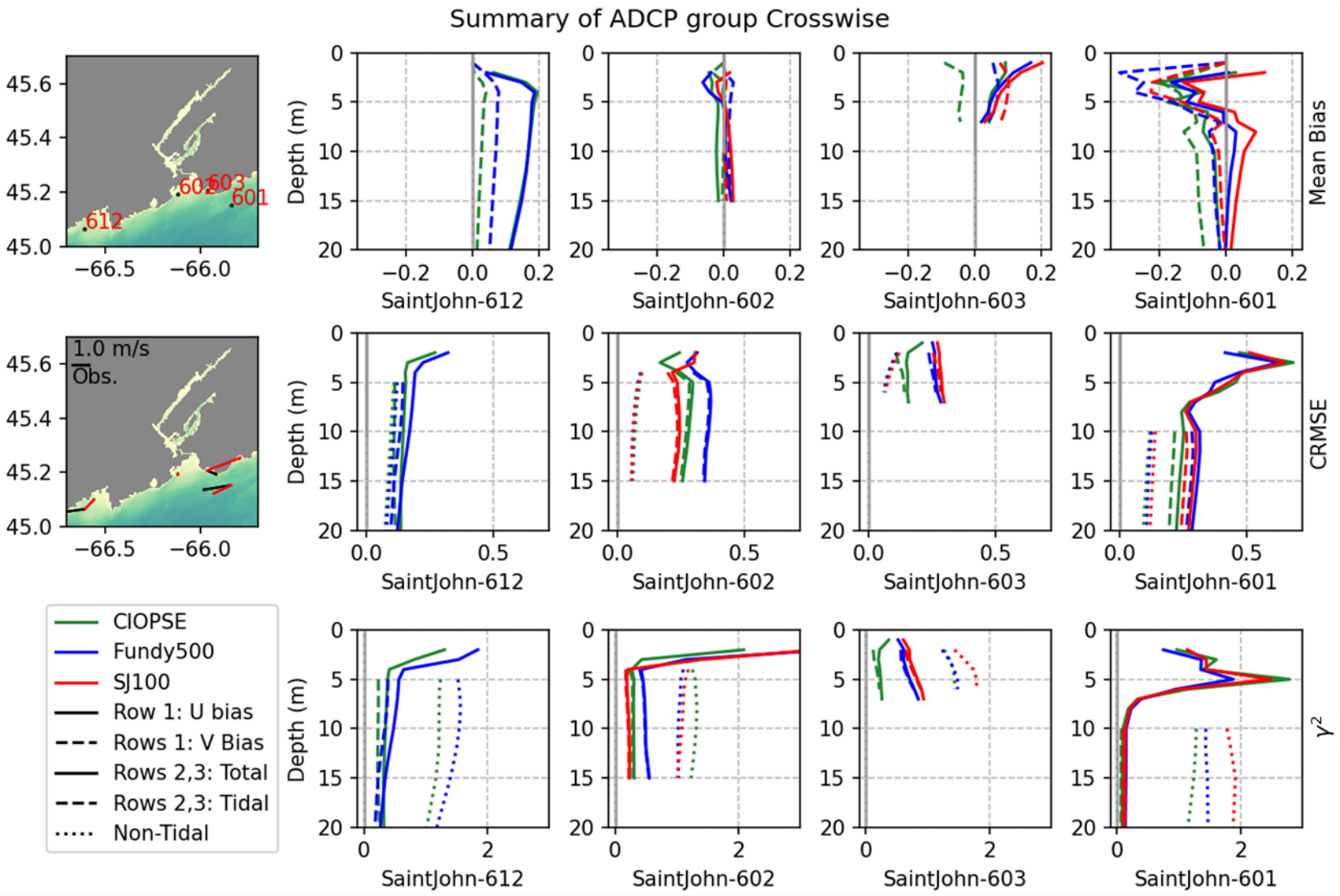

Figure 53: Depth profiles of statistics for a range of moorings in 2017. Top row shows mean bias for u (solid) and v (dashed) velocities. Middle and bottom row show CRMSE and γ^2 for total (solid), tidal (dashed), and non-tidal (dotted) circulation.

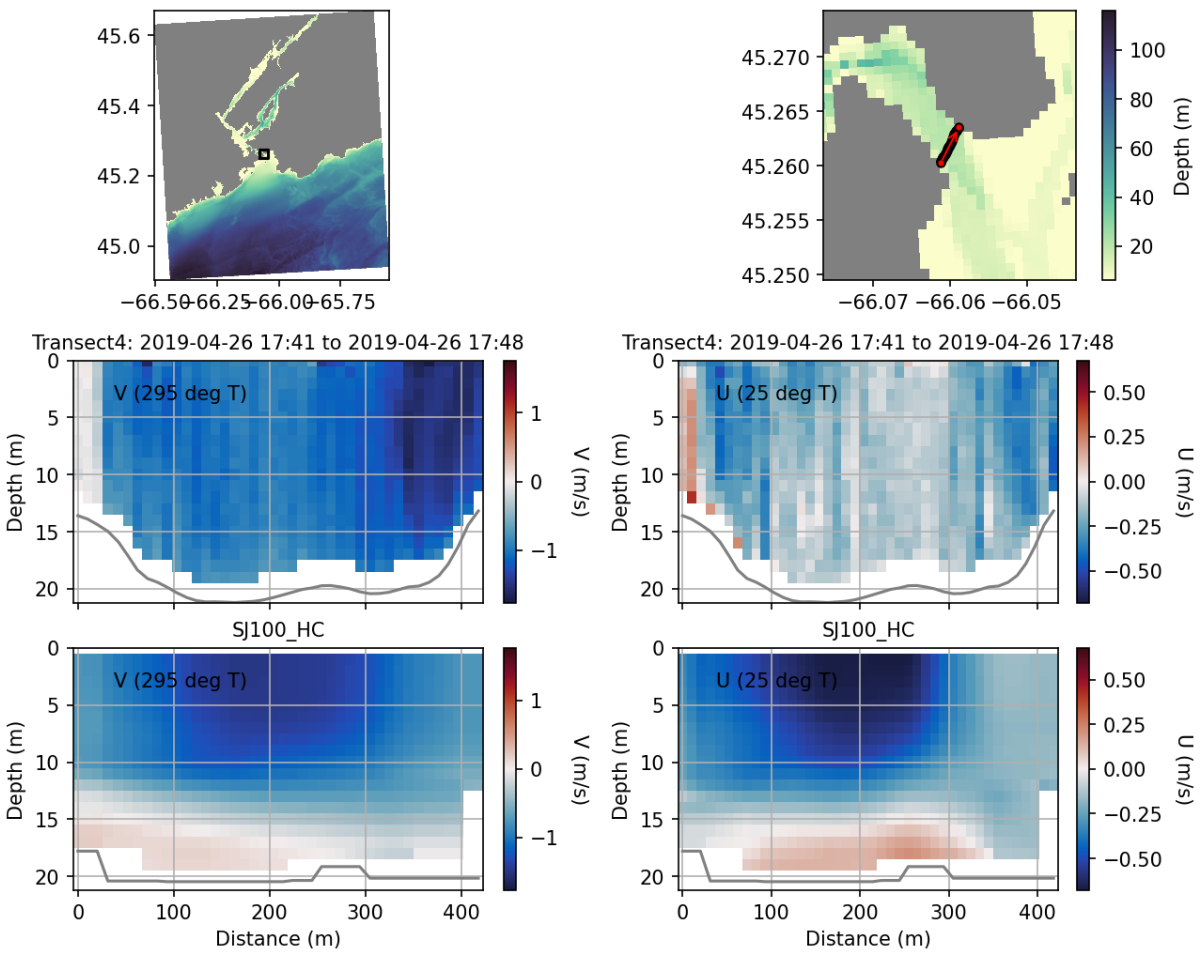

Figure 54: ADCP transect across the mouth of the Saint John river for SJ100 model results. Velocity V is cross-transect (i.e. along-channel), and U is along-transect (i.e. cross-channel). Note the variation of along-transect (cross-channel) velocity. Other models do not adequately represent the transect and so are not shown. Positive flow is out of the river.

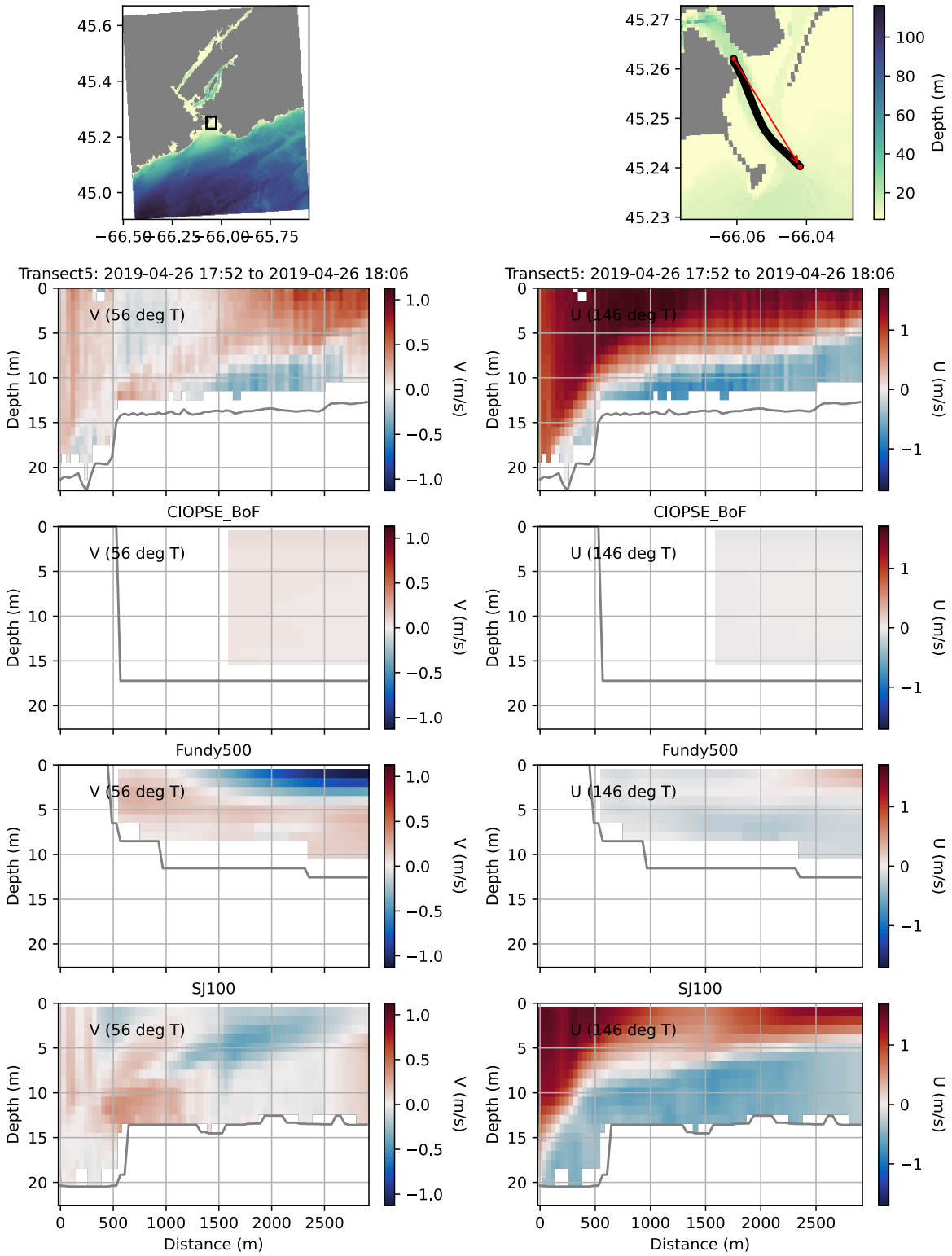

Figure 55: ADCP transect from the river mouth to just past Partridge Island, roughly following the main dredge channel.

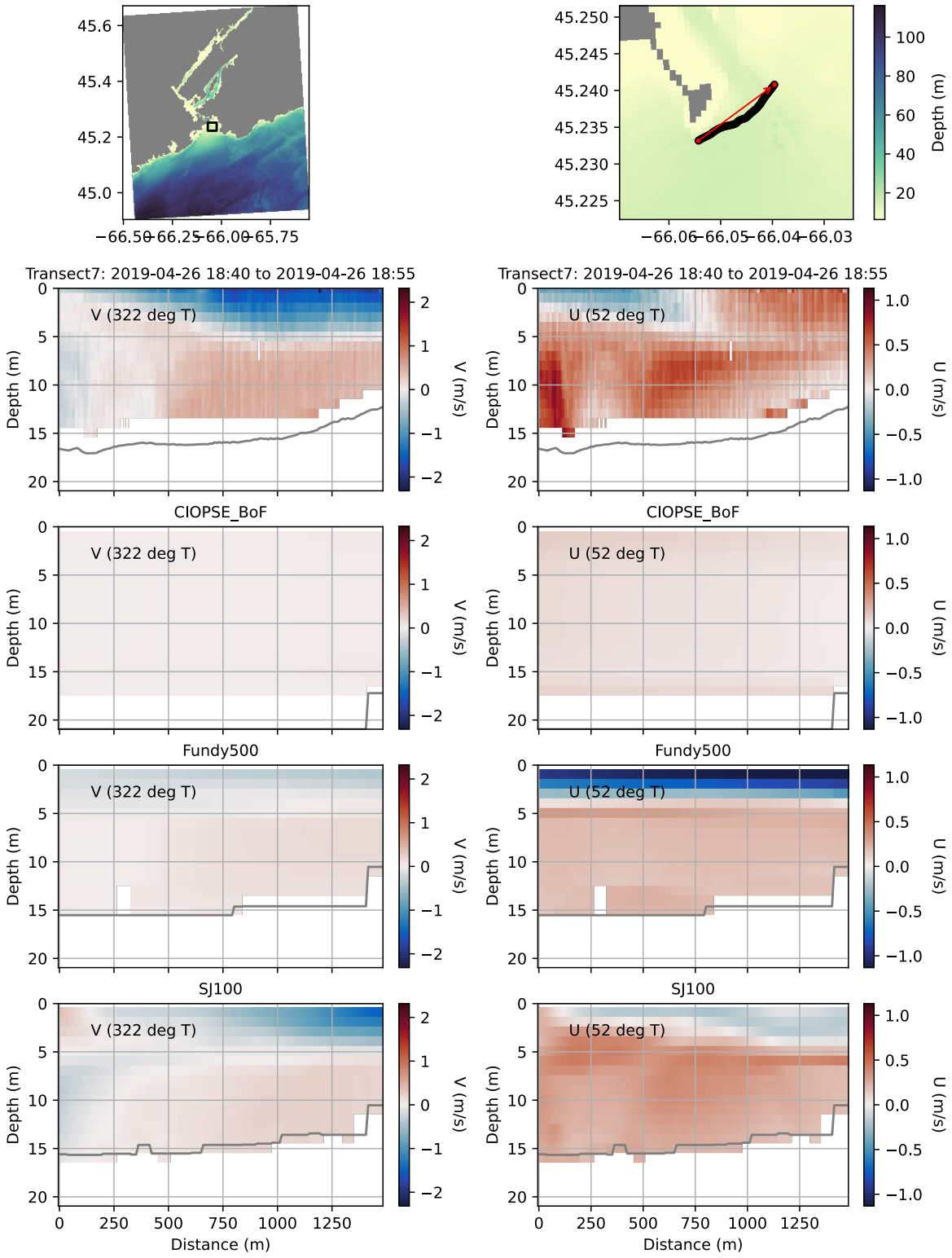

Figure 56: ADCP transect near Partridge Island.

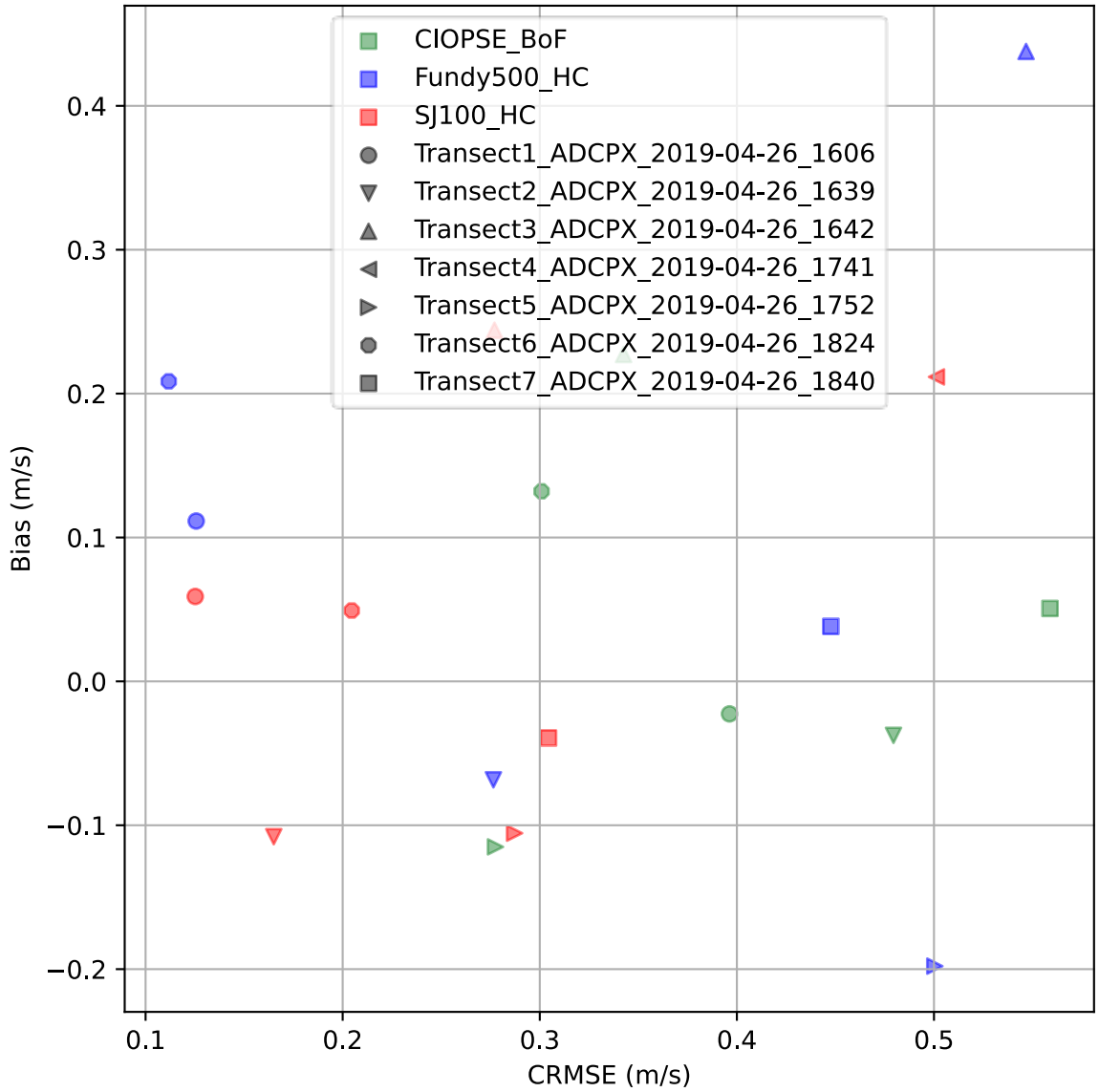

Figure 57: Bias vs CRMSE for all seven ADCP transects for all three models.

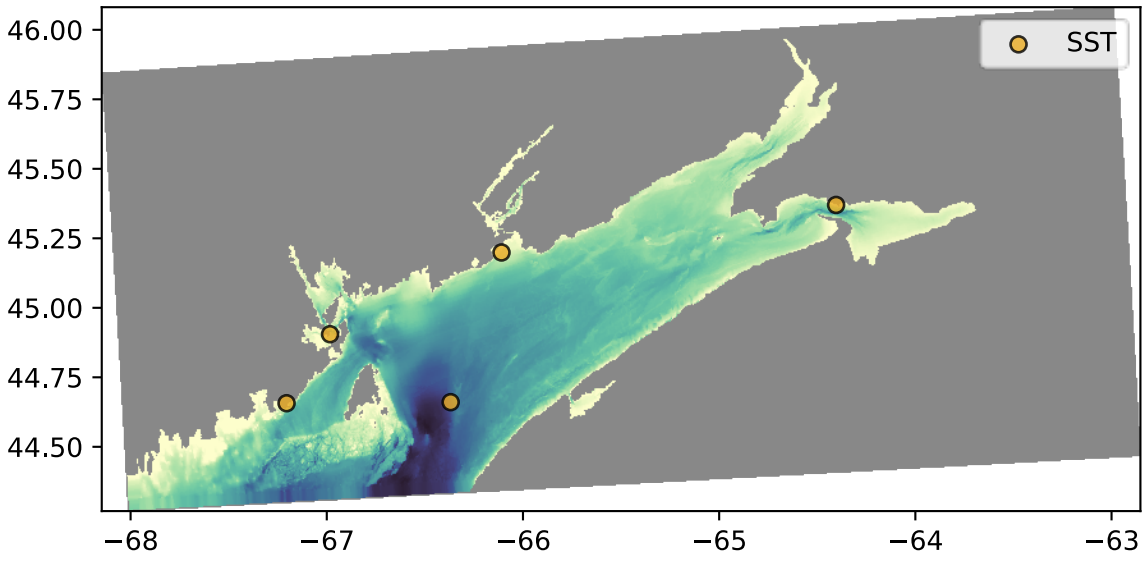

Figure 58: Locations of SST observations.

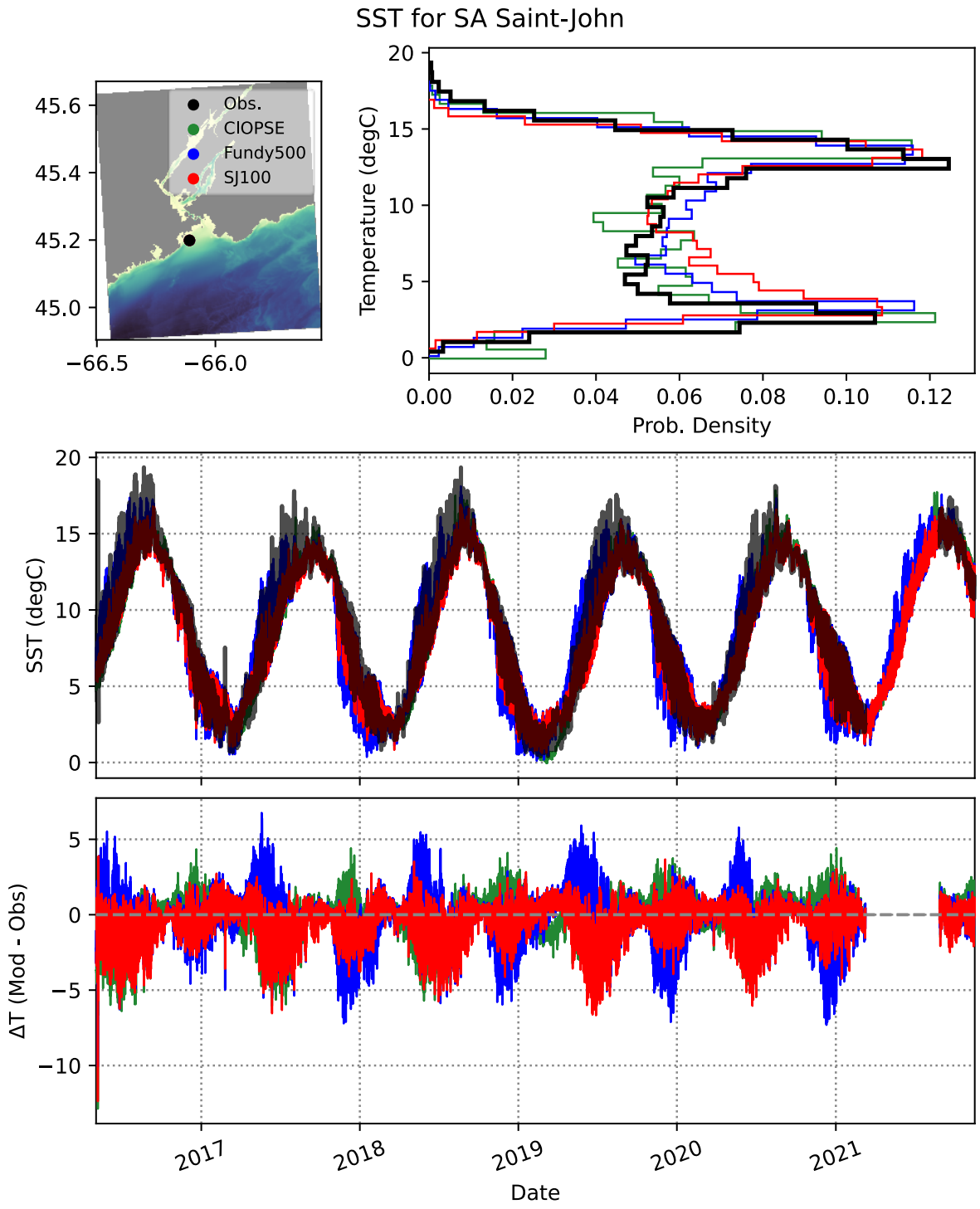

Figure 59: SST at the SmartAtlantic buoy outside of Saint John harbour. Time series covers the full span of the hindcast.

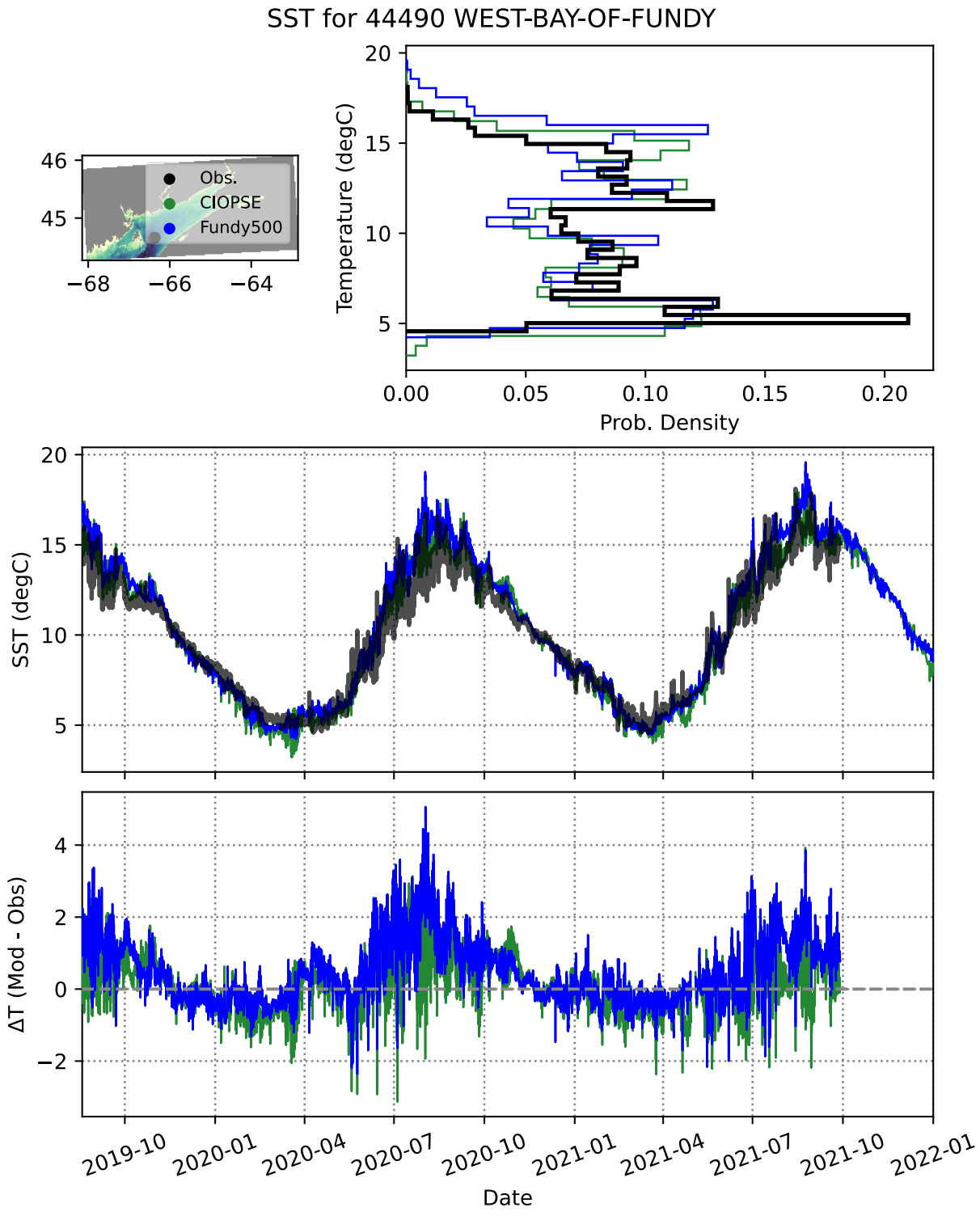

Figure 60: SST at the buoy 44490, between Grand Manan and Digby Neck.

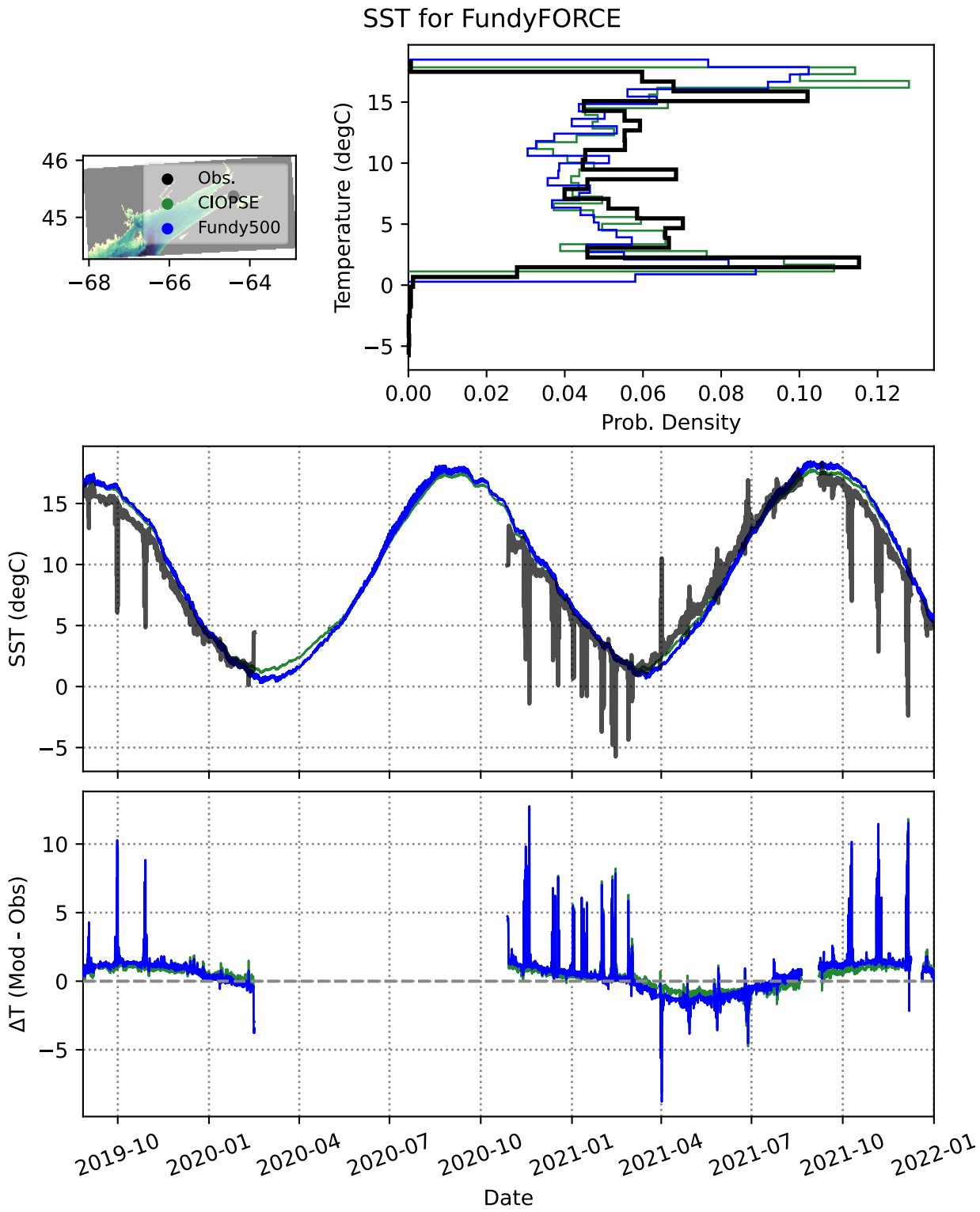

Figure 61: SST at Fundy FORCE, near Parrsboro, NS. Cold spikes may be due to the buoy sloshing and tipping into the air, rather than an accurate SST measurement.

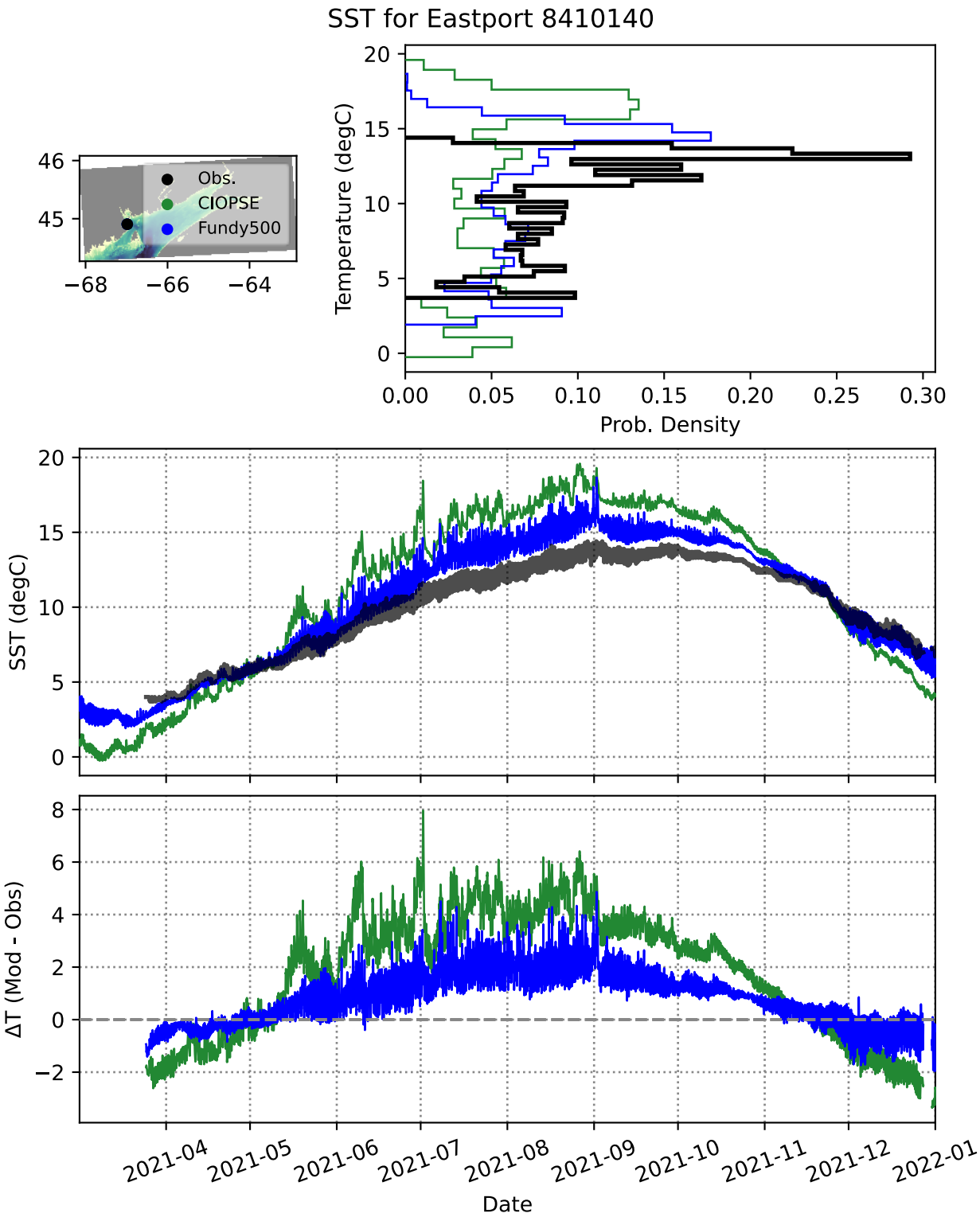

Figure 62: SST at Eastport, Maine.

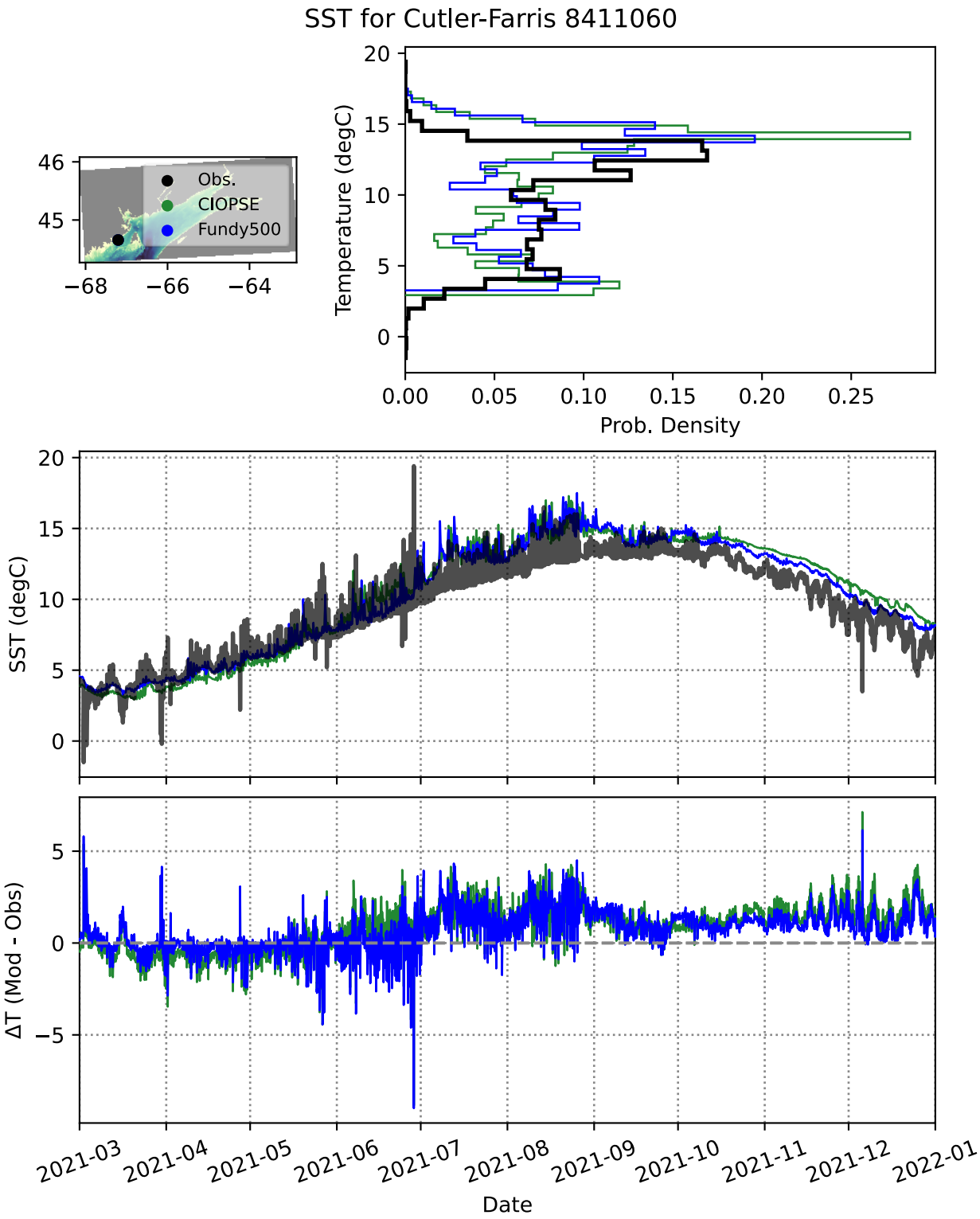

Figure 63: SST at Cutler Farris, Maine.

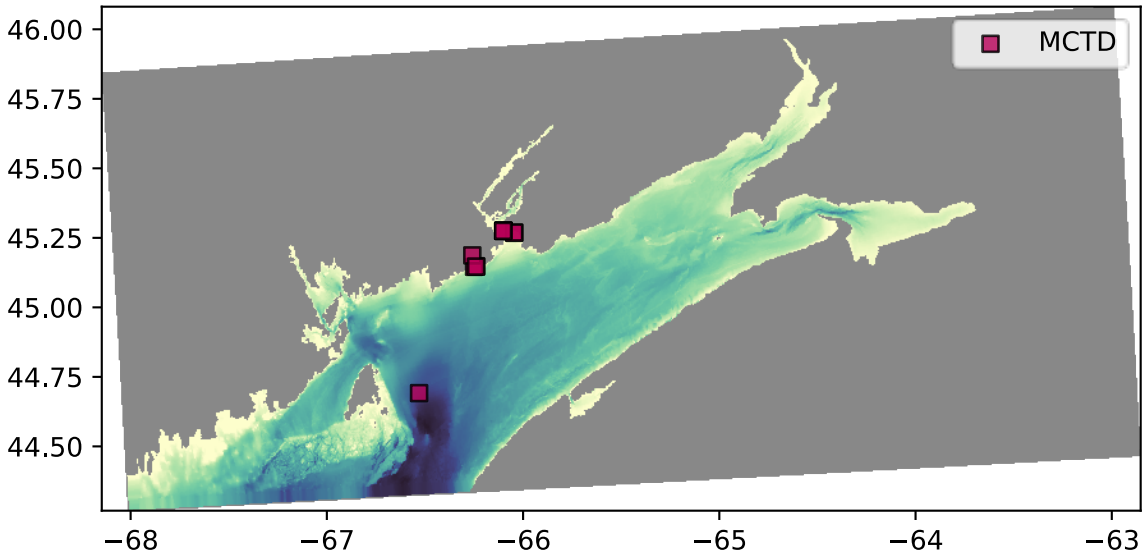

Figure 64: Locations of moored CTDs.

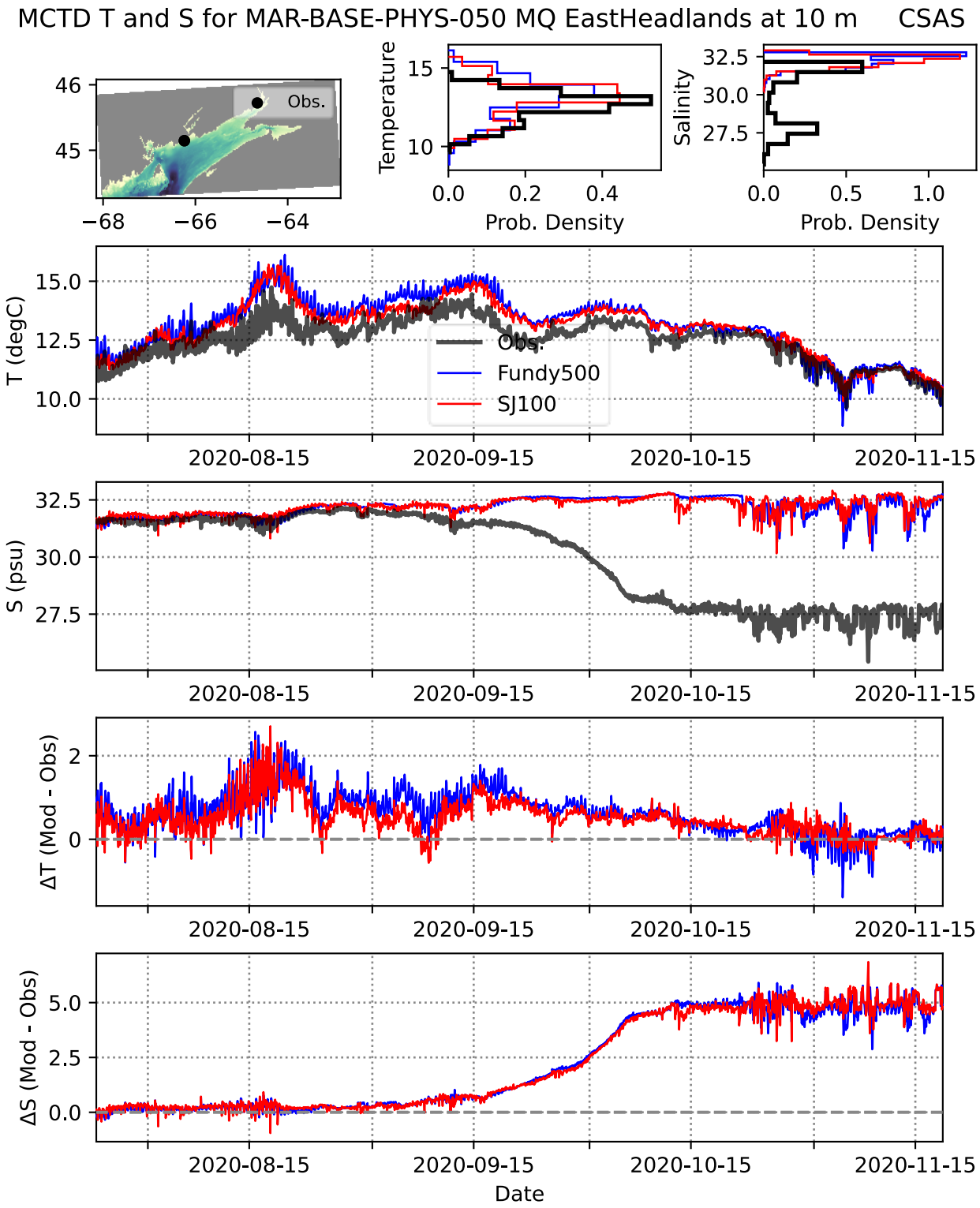

Figure 65: T and S near the seabed at Musquash eastern headlands. Observed salinity drops mid-series likely due to pulse of fresh water from the Musquash river.

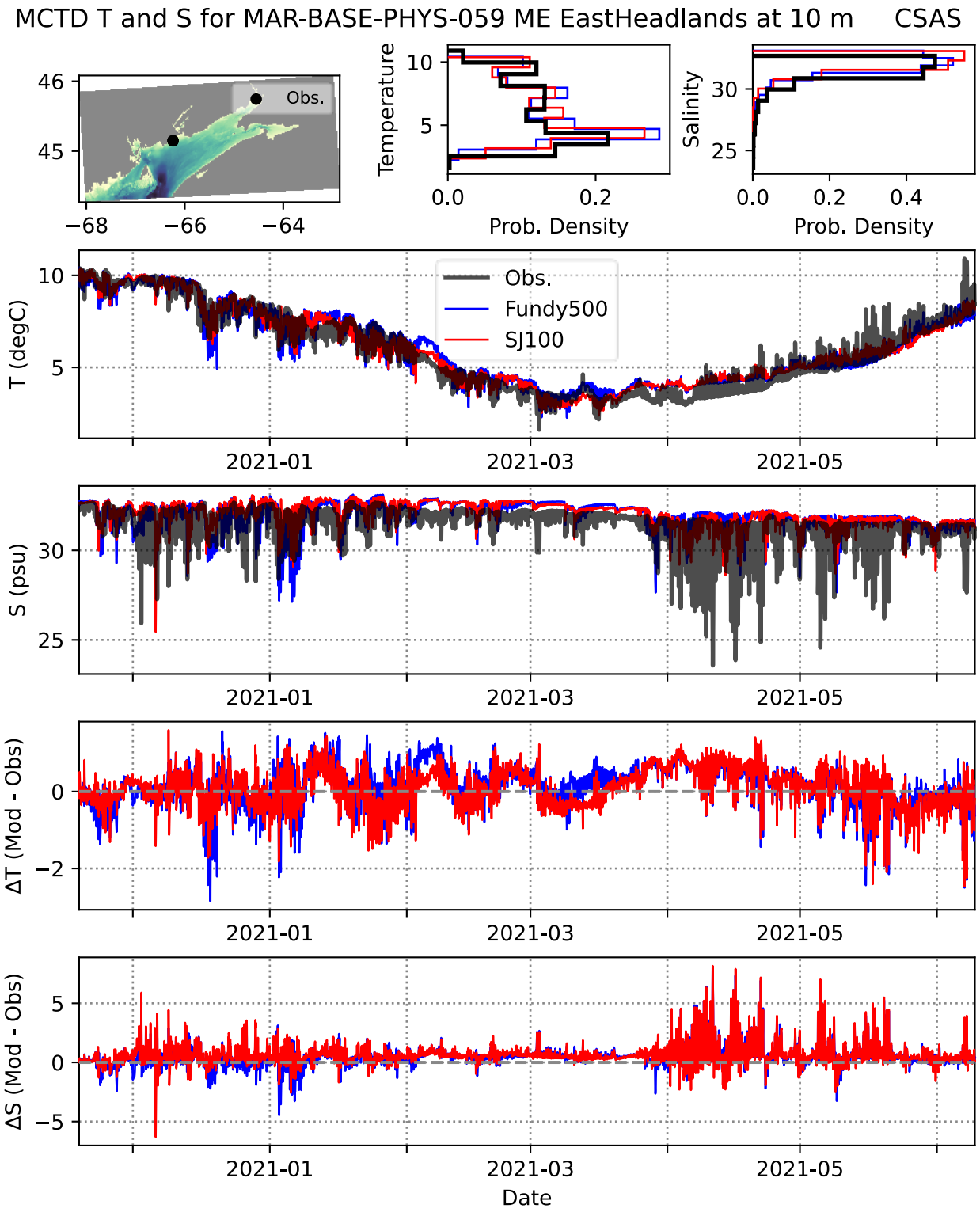

Figure 66: T and S near the seabed at Musquash eastern headlands.

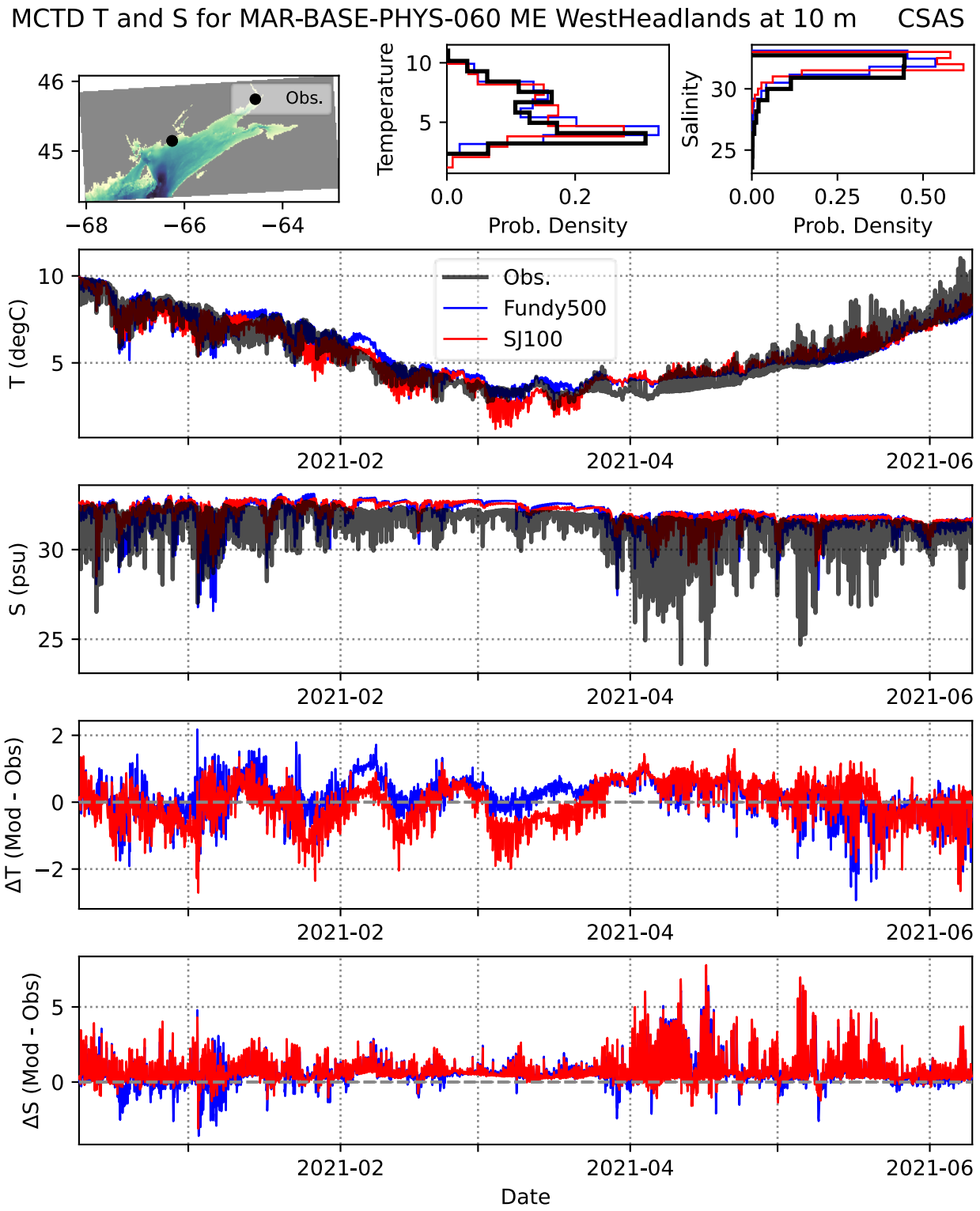

Figure 67: T and S near the seabed at Musquash western headlands.

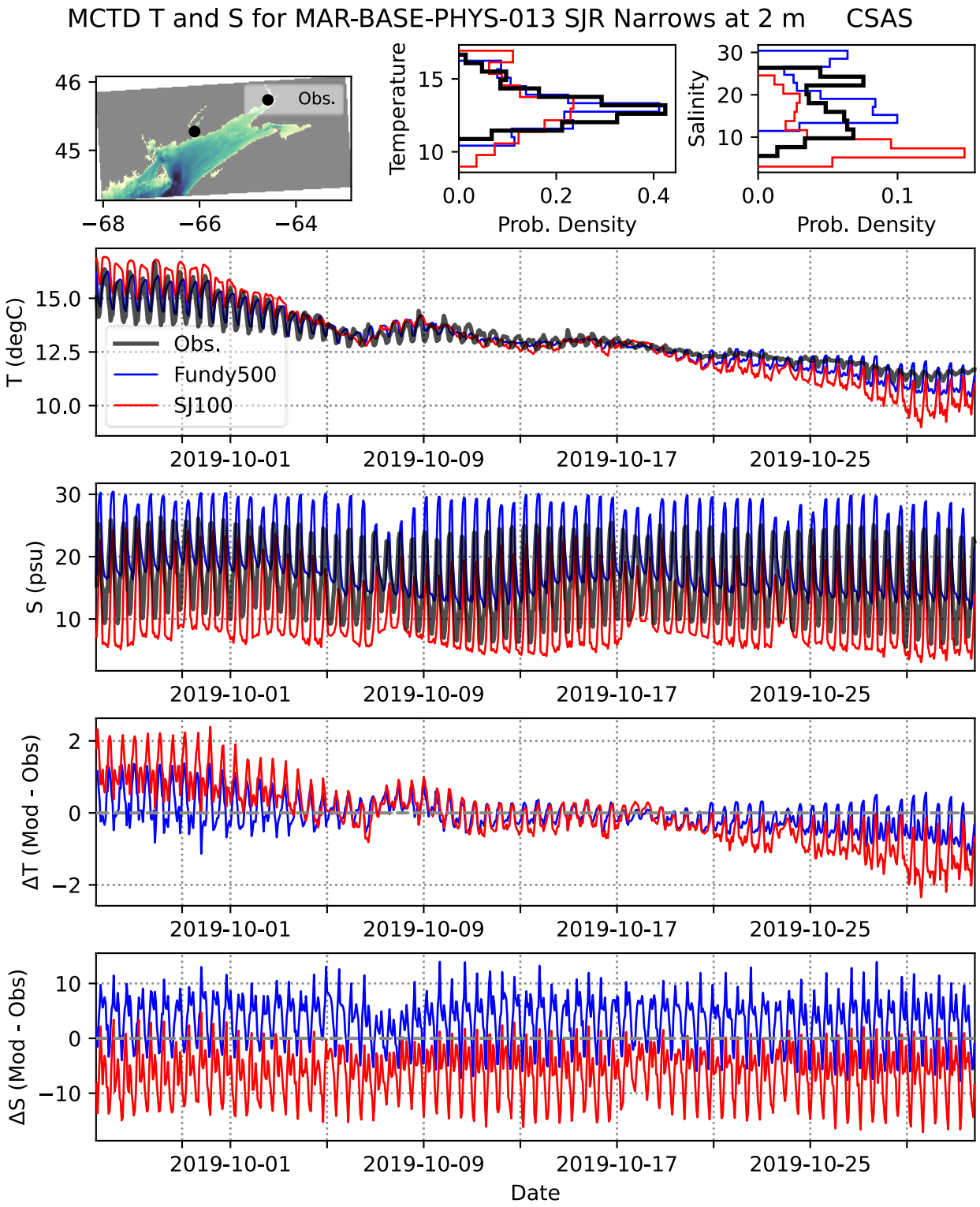

Figure 68: Near surface T and S just upstream of Reversing Falls.

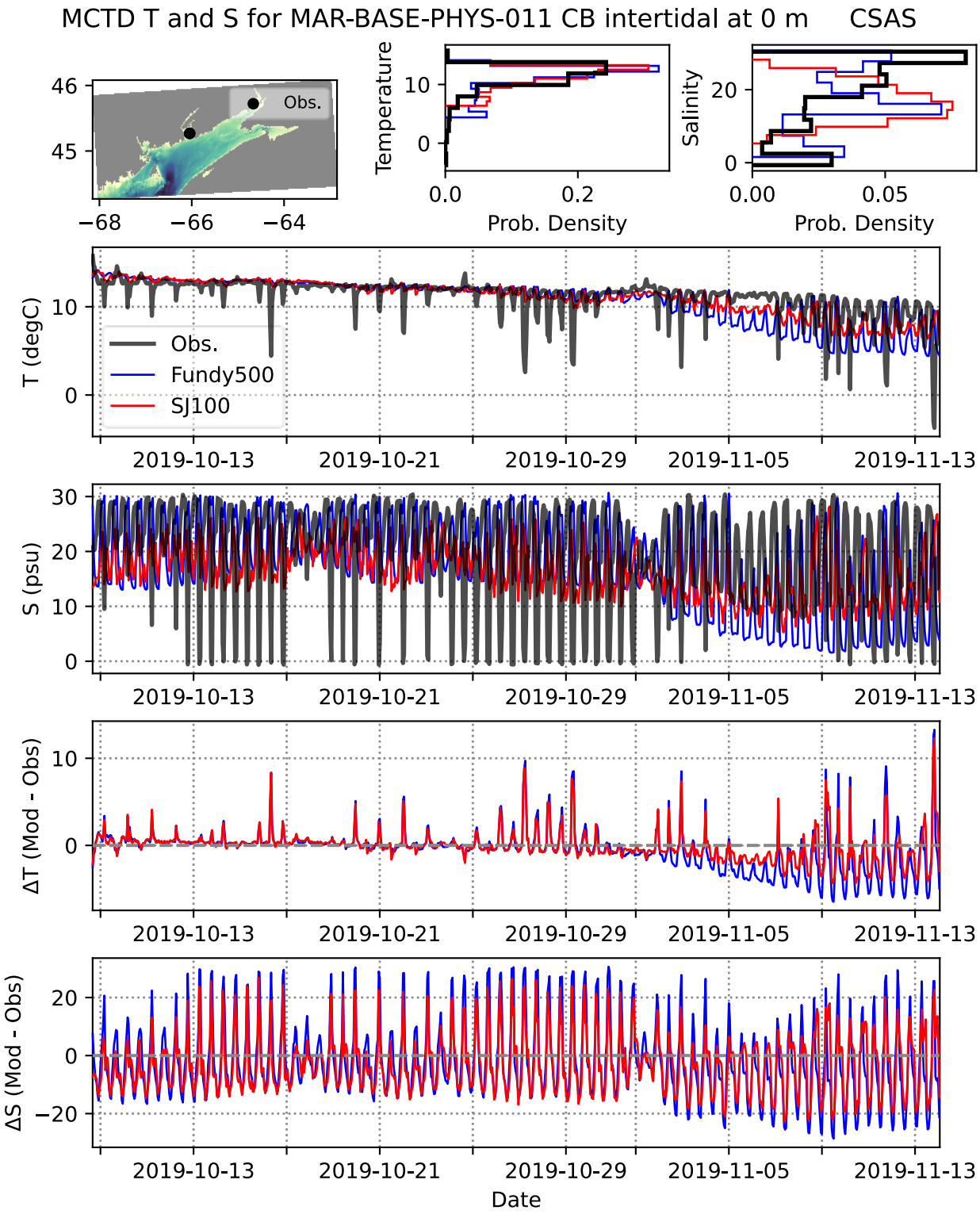

Figure 69: Surface T and S in the intertidal zone of Courtenay Bay. This area is not well represented in the model but is included to assess how well the models capture the low frequency signal.

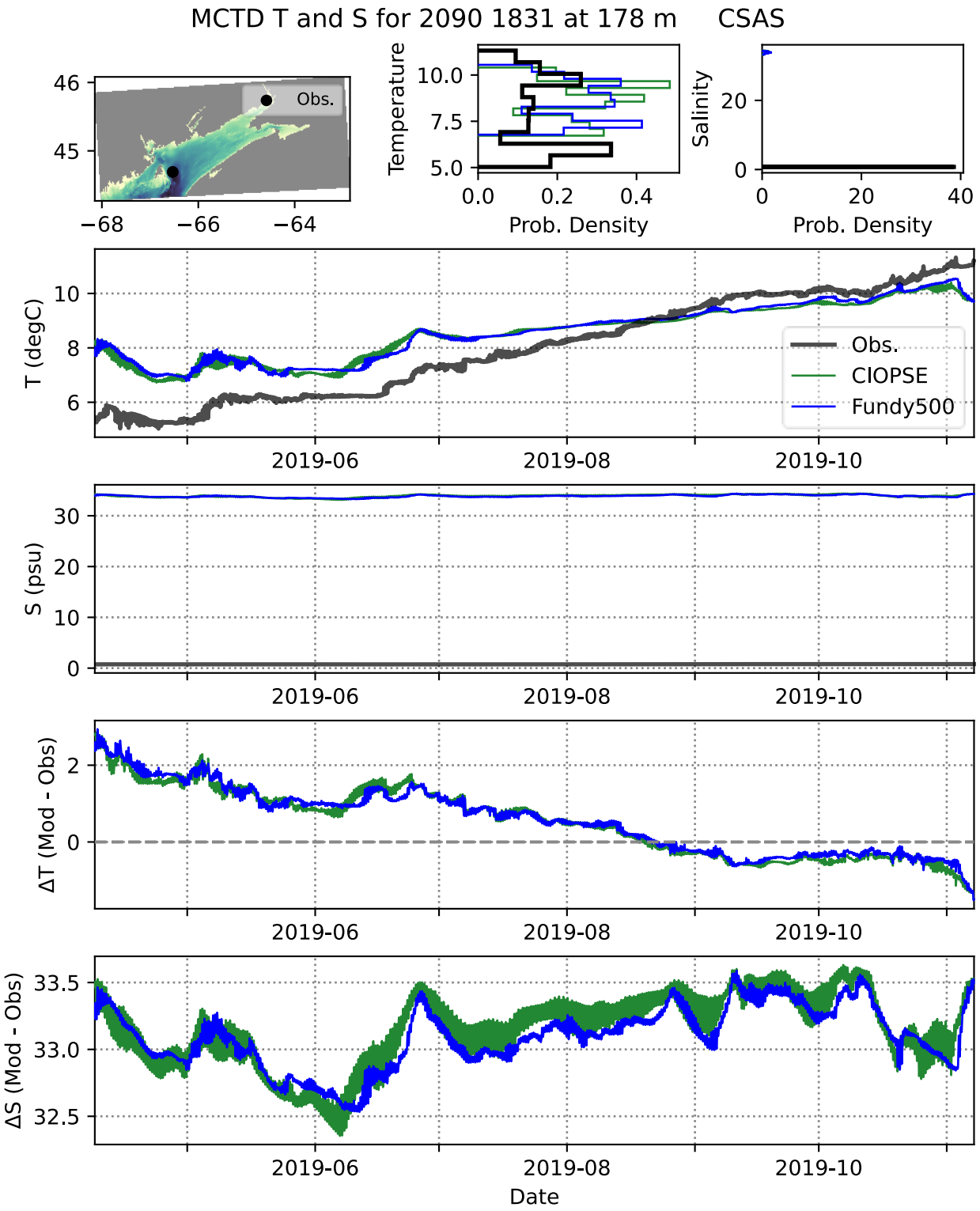

Figure 70: T and S at depth off Grand Manan. No observed salinity data was available but is included to be able to compare the characteristics of the CIOPS-E and Fundy500 time series.

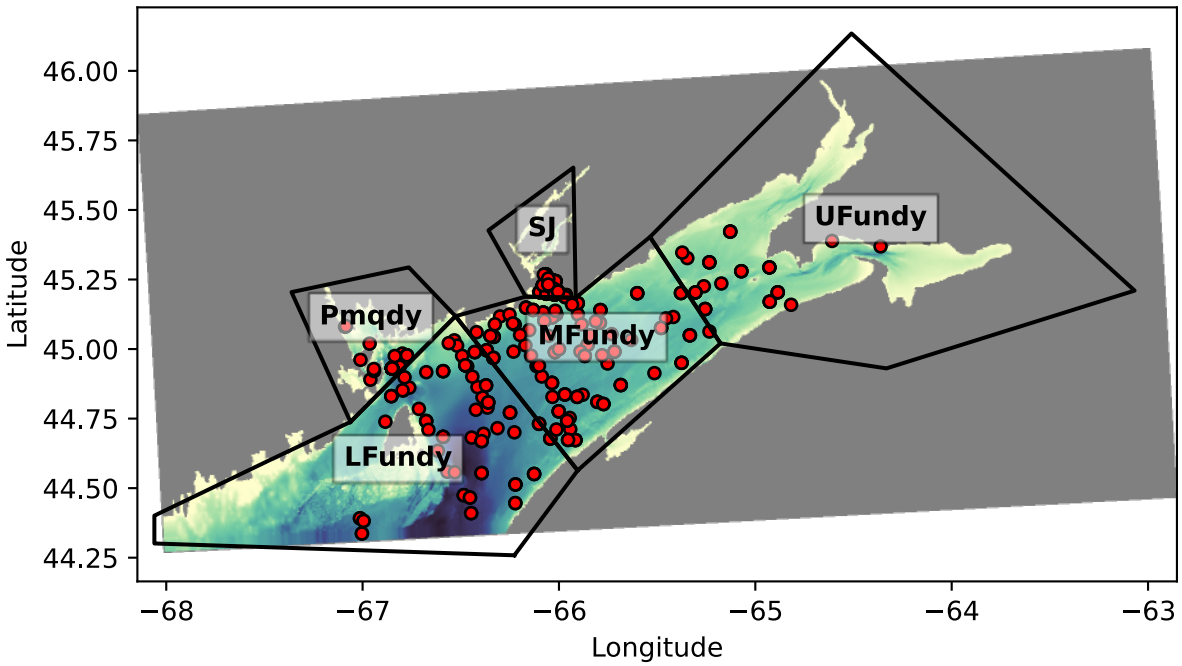

Figure 71: Regions of port model domain used for CTD aggregate statistics. Red dots indicate CTD casts used, spanning the entire hindcast.

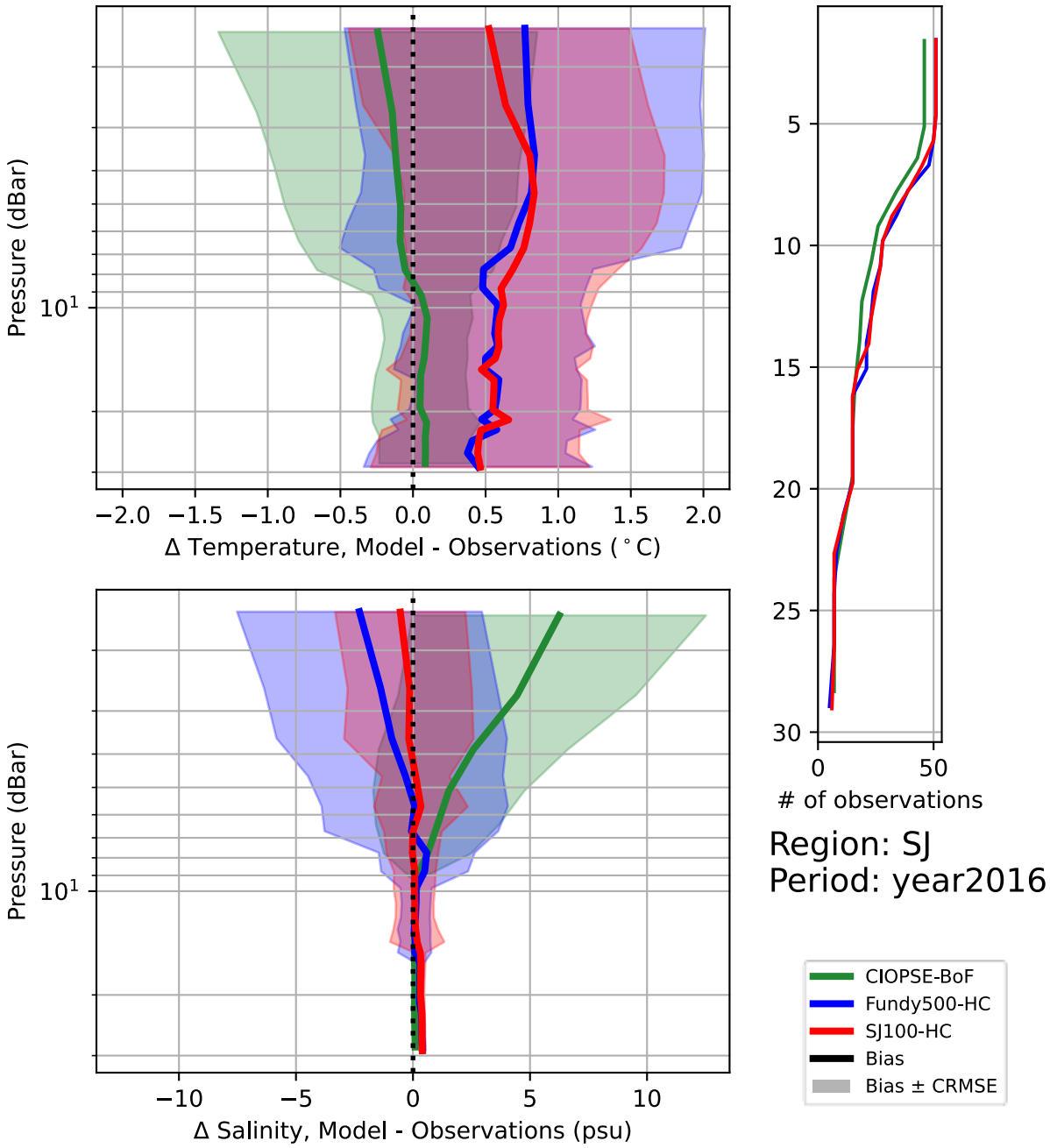

Figure 72: Depth profile of T and S bias and CRMSE for the Saint John (SJ) region for 2016.

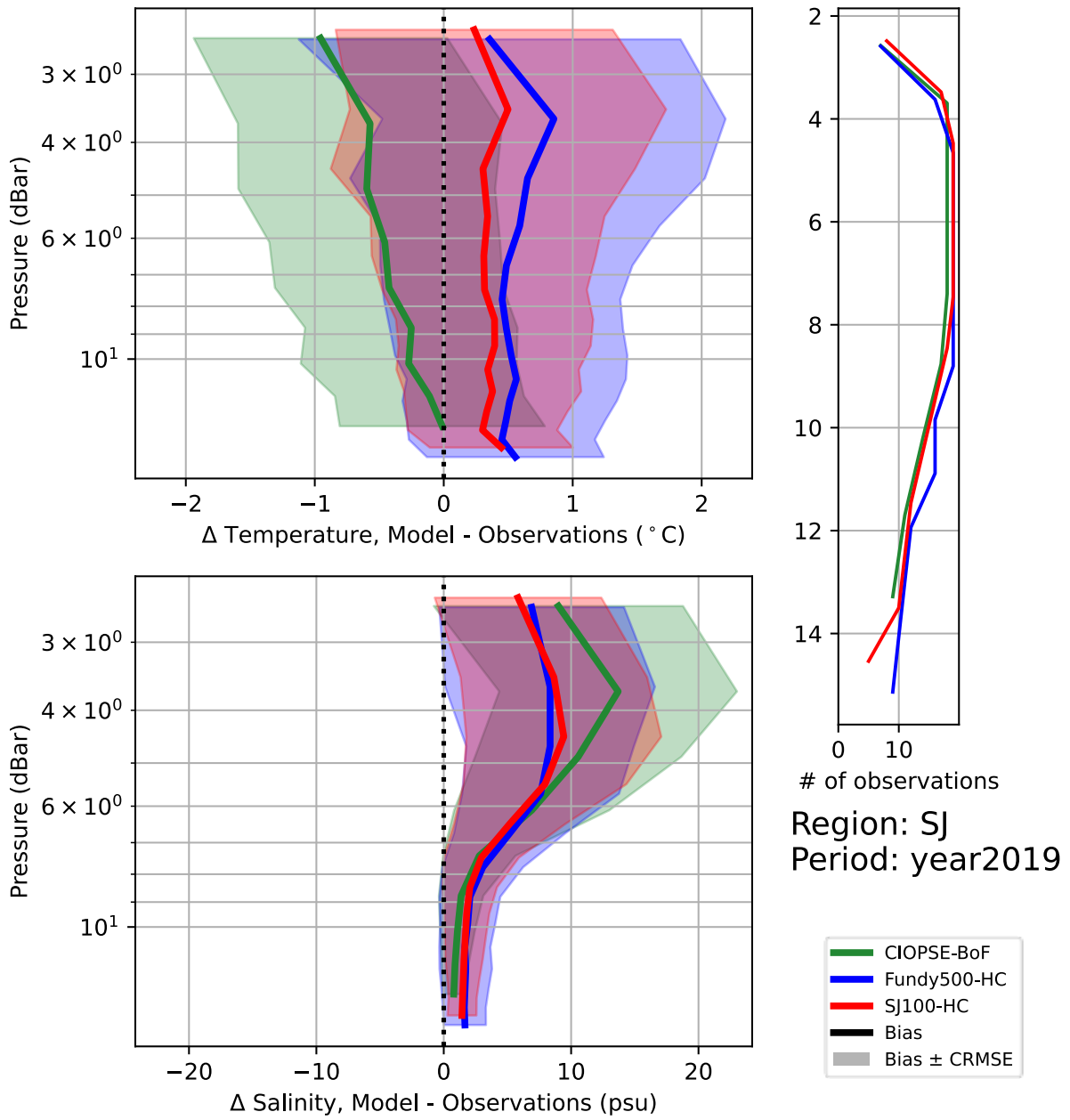

Figure 73: Depth profile of T and S bias and CRMSE for the Saint John (SJ) region for 2019. The bulk of the profiles from 2019 were taken during the spring freshet and flood, and the models are salty relative to the unusually fresh near-surface depths.

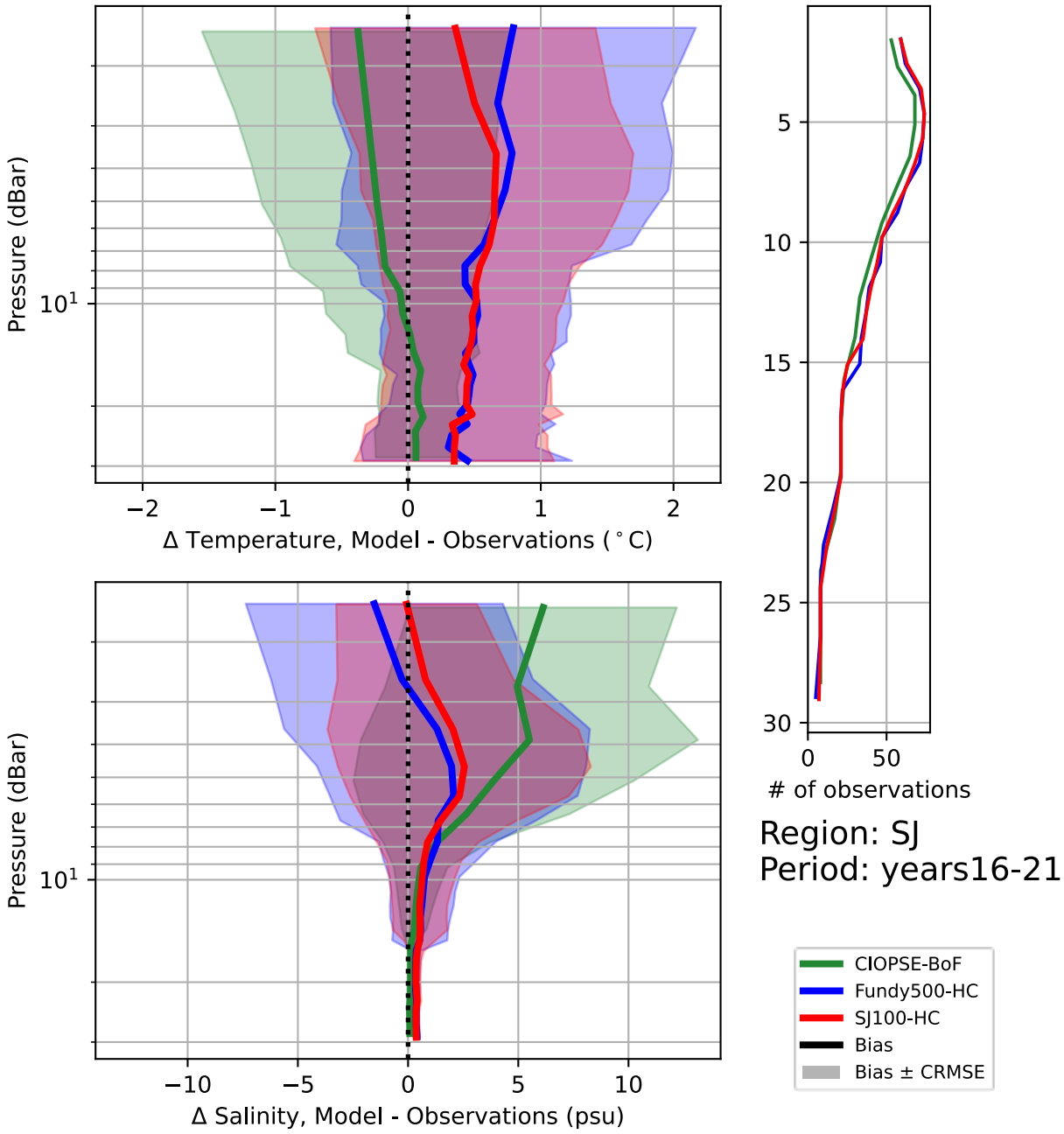

Figure 74: Depth profile of T and S bias and CRMSE for the Saint John (SJ) region for the entire hindcast.

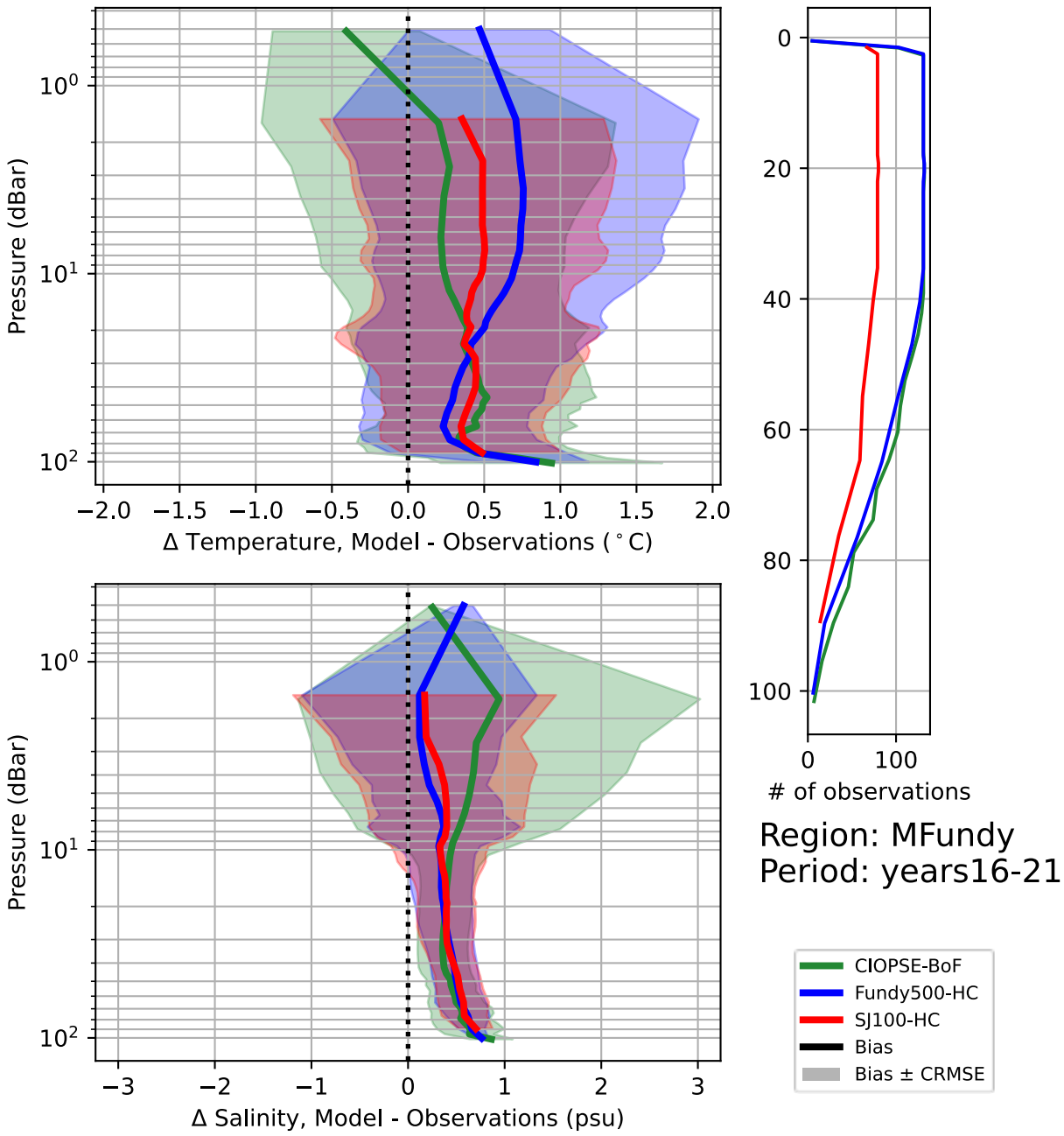

Figure 75: Depth profile of T and S bias and CRMSE for the Mid-Fundy (MFundy) region.

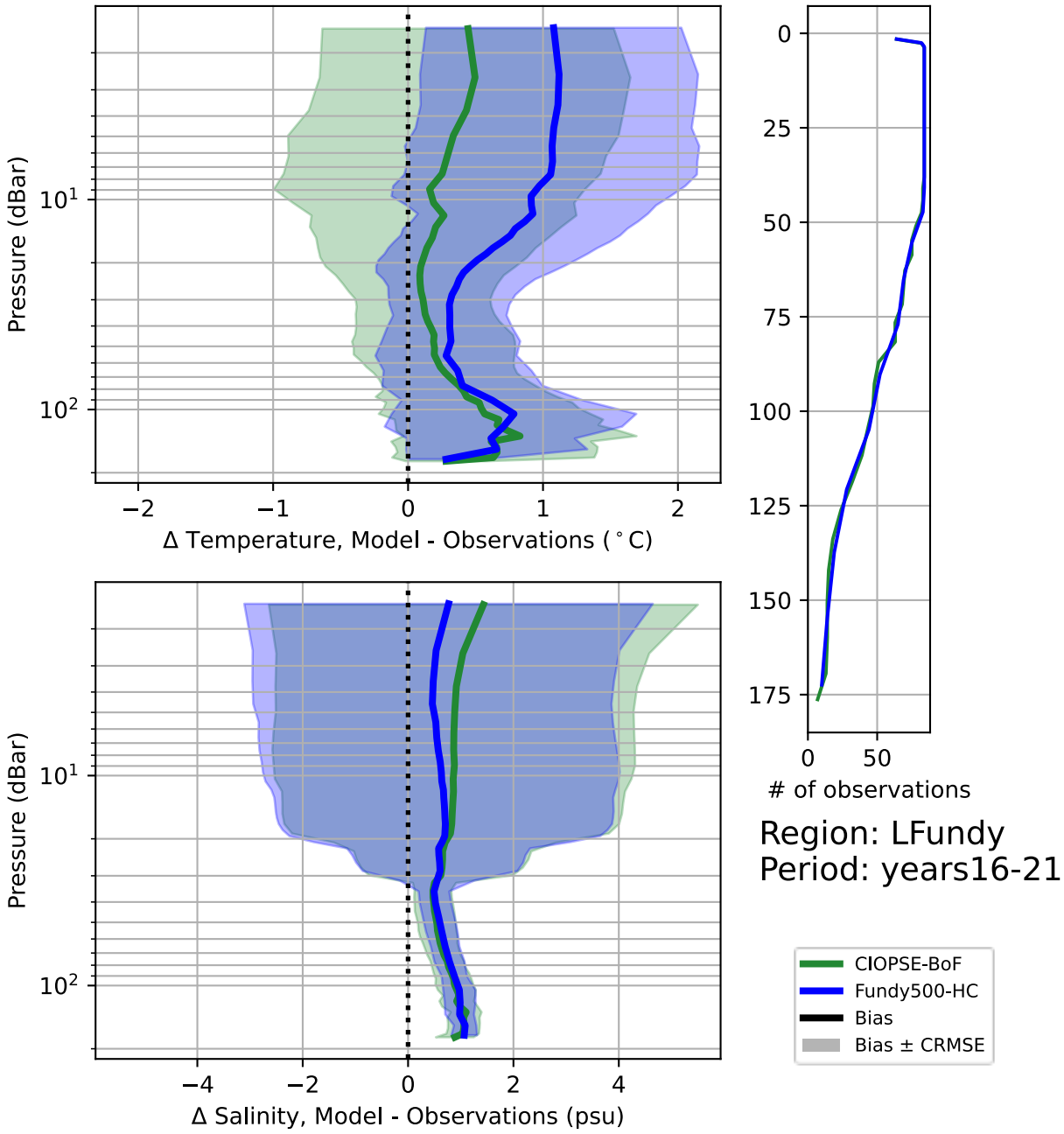

Figure 76: Depth profile of T and S bias and CRMSE for the Lower Fundy (LFundy) region.

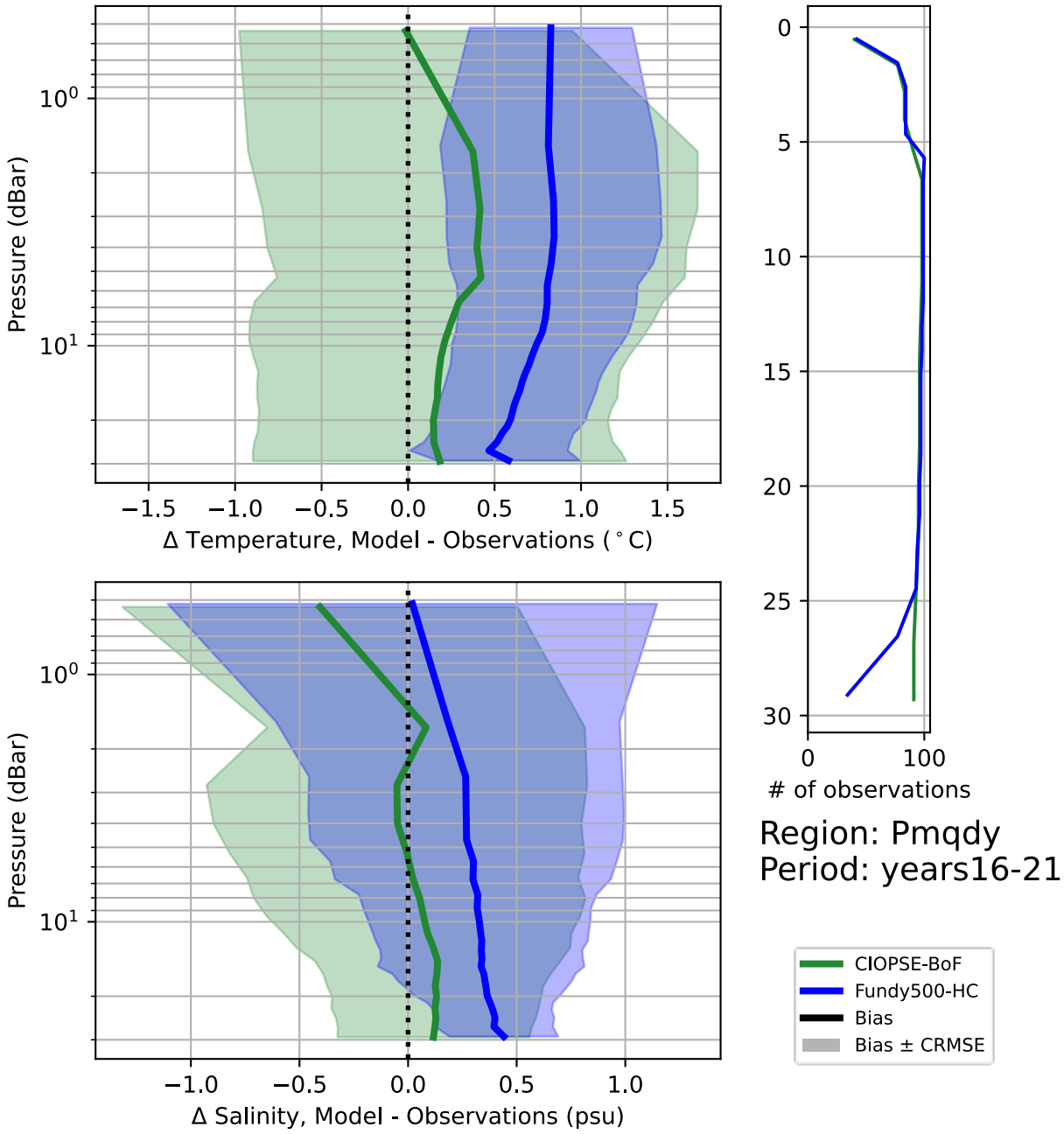

Figure 77: Depth profile of T and S bias and CRMSE for the Passamaquoddy Bay (Pmqdy) region.

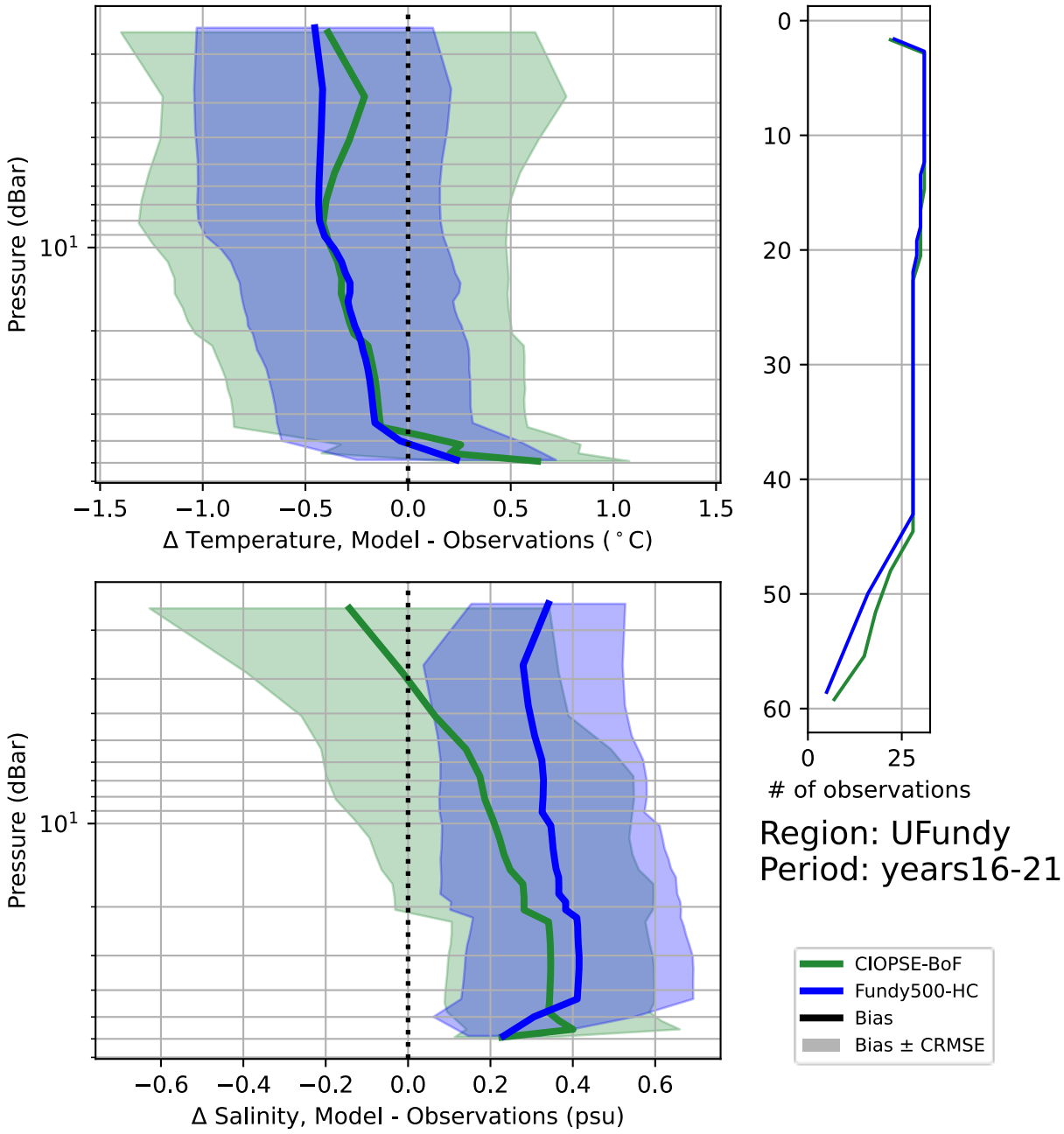

Figure 78: Depth profile of T and S bias and CRMSE for the Upper Fundy (UFundy) region.

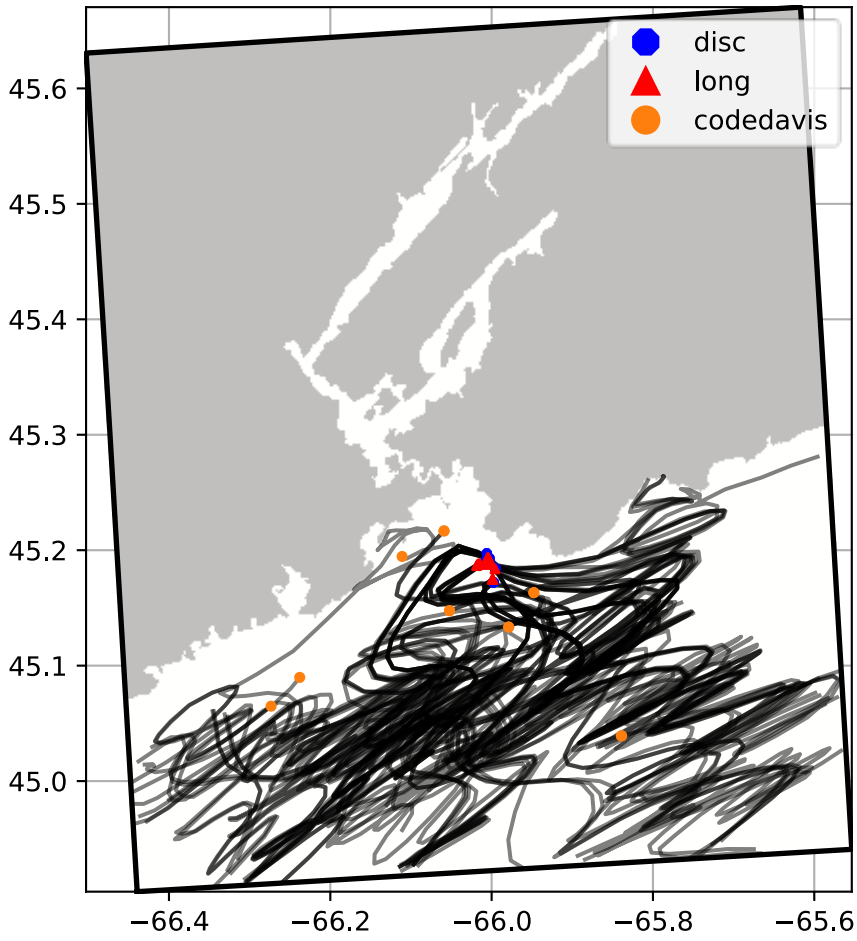

Figure 79: Map of all drift tracks used for SJ100 and Fundy500 evaluation. All tracks are cropped to the SJ100 model domain boundaries.

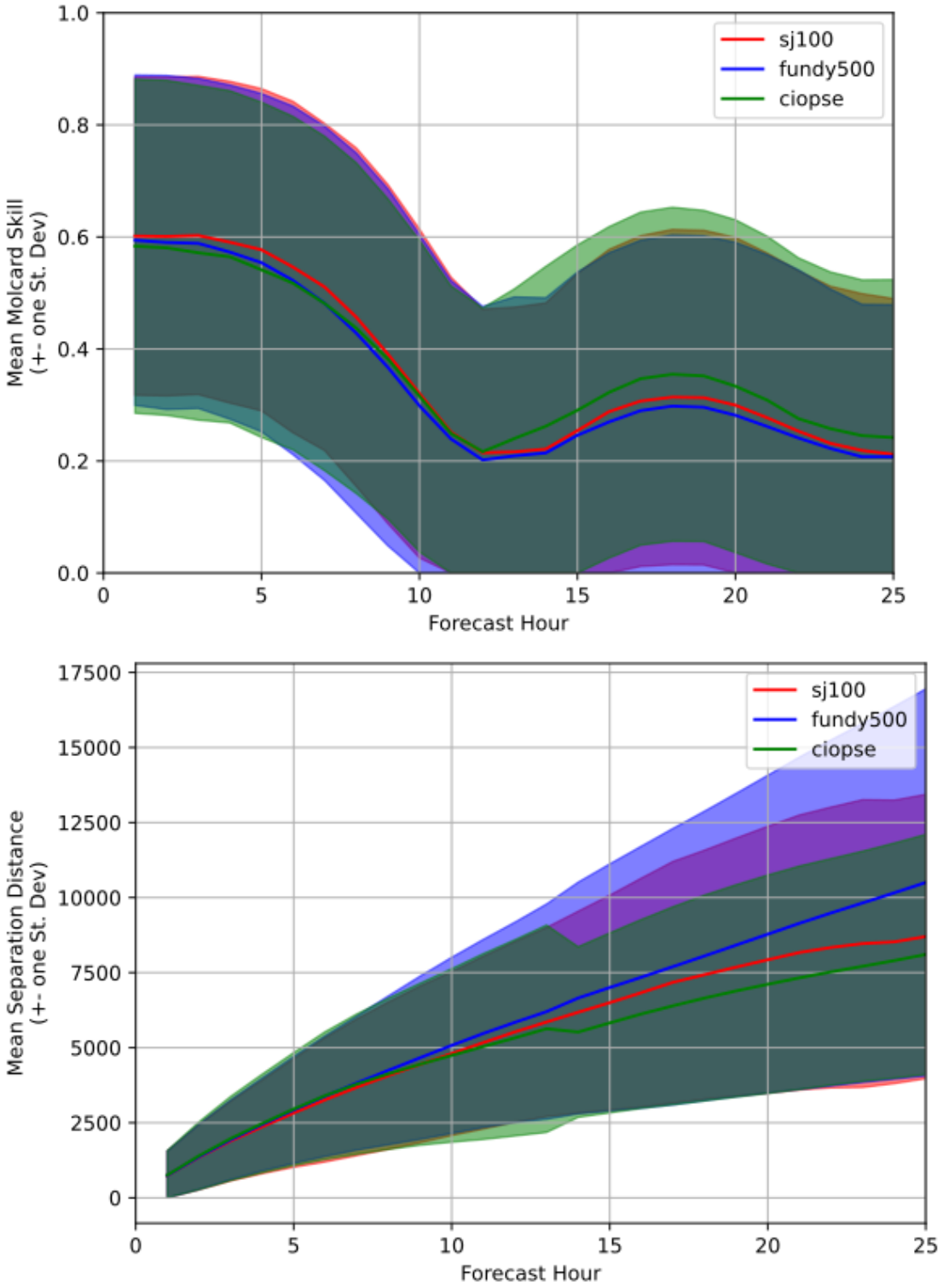

Figure 80: Mean Molcard skill (top) and separation distance (bottom) for Davis drifters for 25 hours following simulated drifter release. Color indicates hydrodynamics model. Shaded region indicates ± 1 standard deviation around the mean.

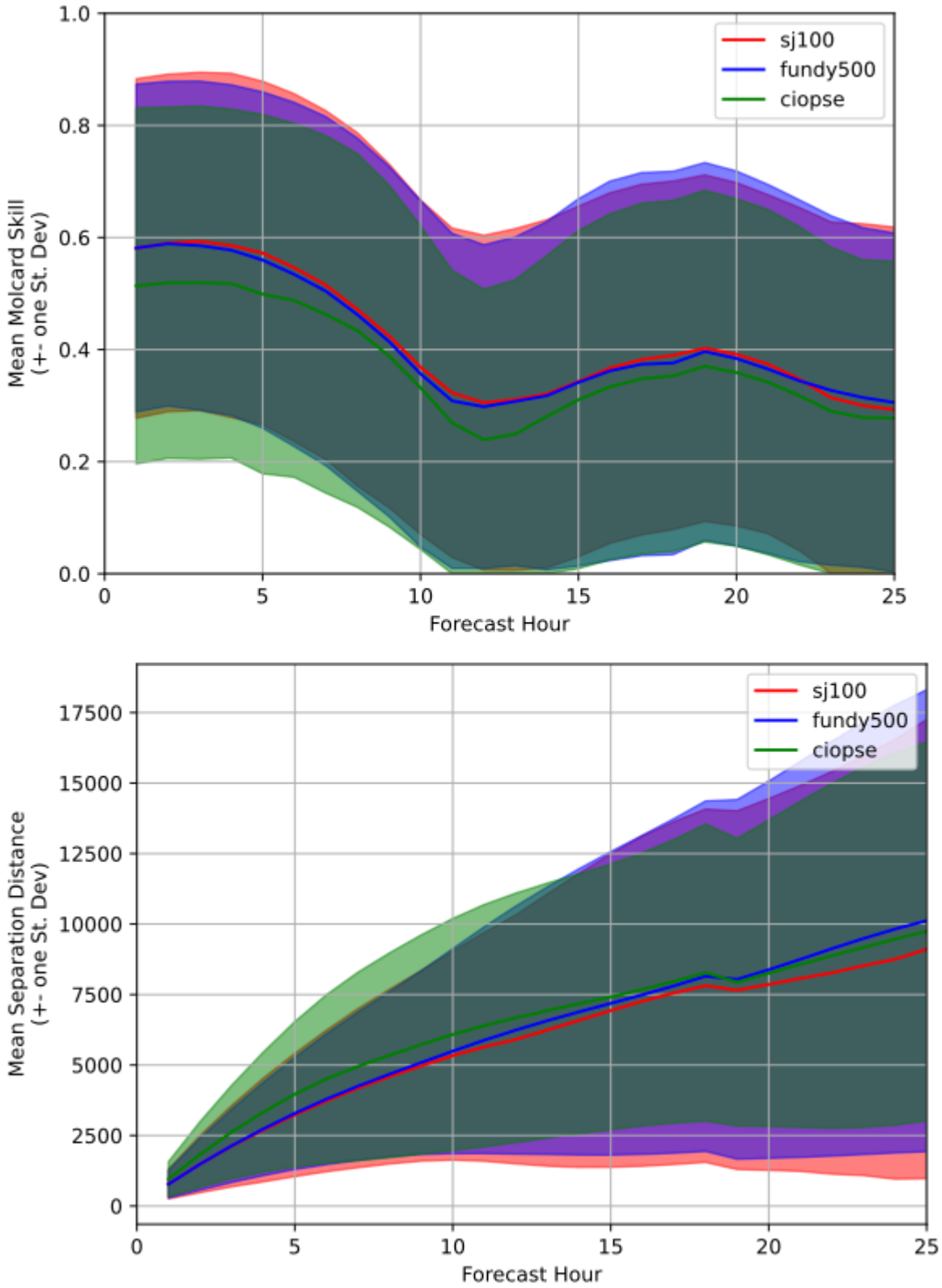

Figure 81: Mean Molcard skill (top) and separation distance (bottom) for SABS Long drifters for 25 hours following simulated drifter release. Color indicates hydrodynamics model. Shaded region indicates ± 1 standard deviation around the mean.

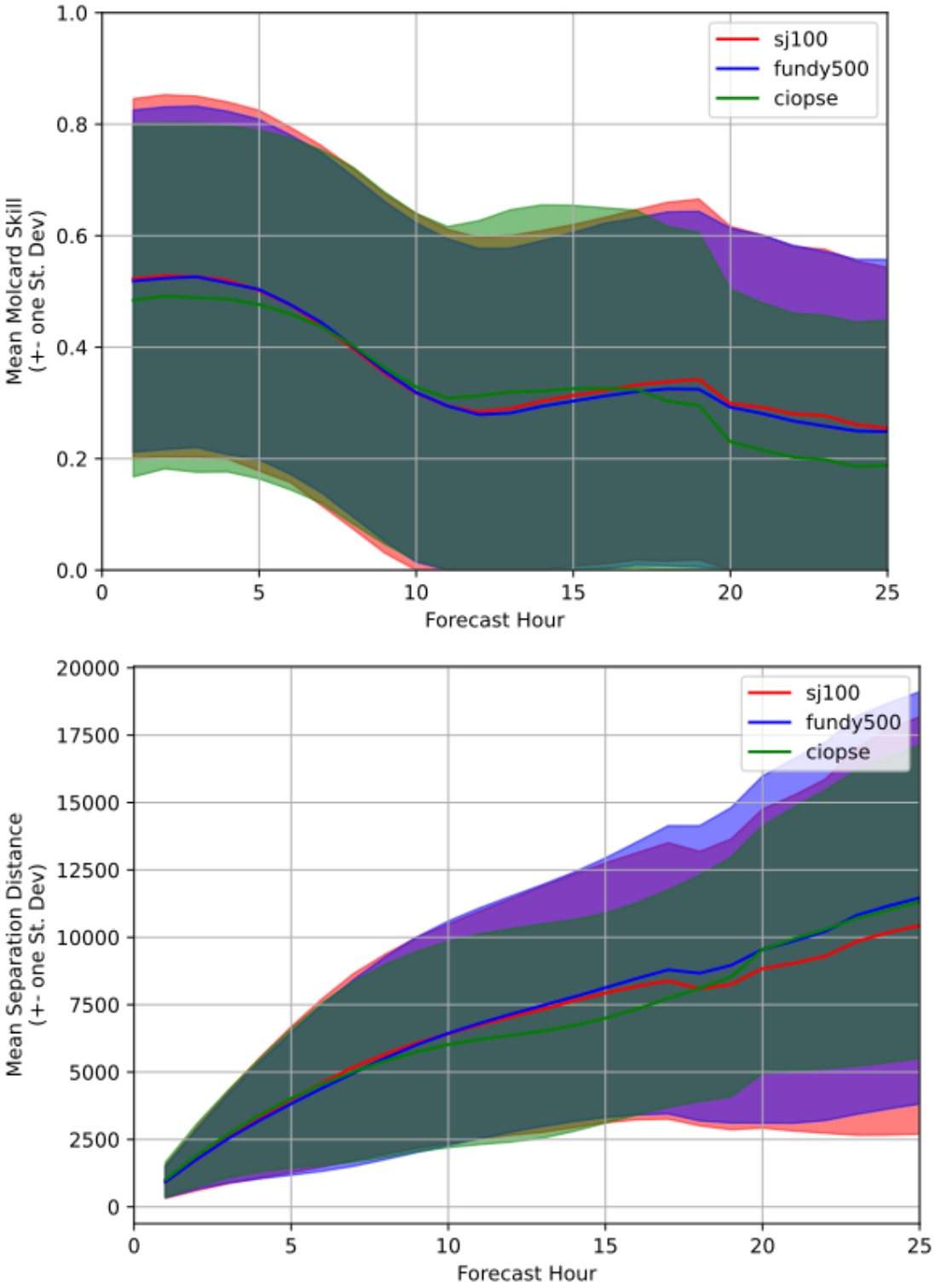

Figure 82: Mean Molcard skill (top) and separation distance (bottom) for SABS Disk drifters for 25 hours following simulated drifter release. Color indicates hydrodynamics model. Shaded region indicates ± 1 standard deviation around the mean.

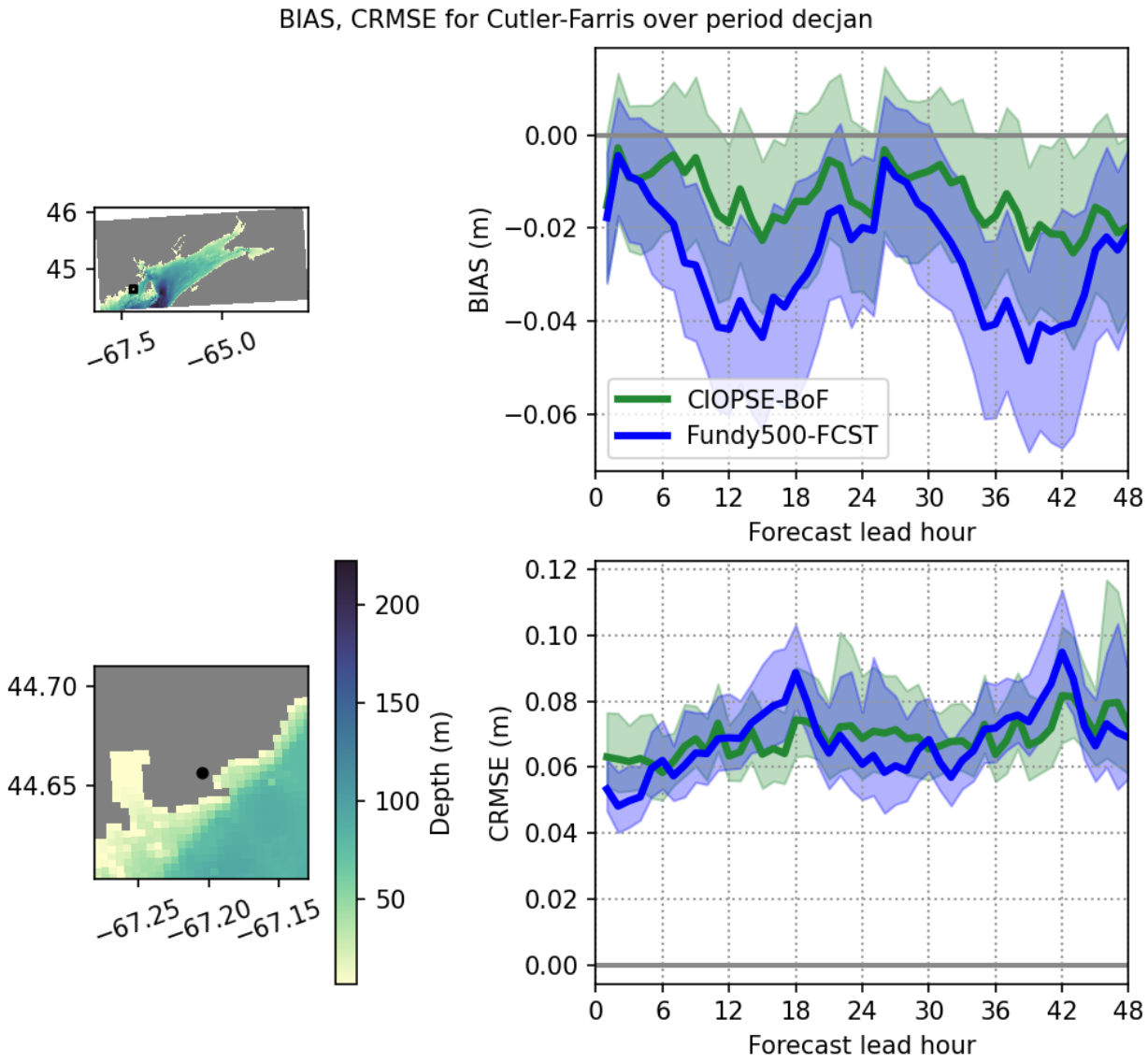

Figure 83: Non-tidal water level bias and CRMSE as a function of lead at Cutler Farris for the batch of 48-hour long 00Z forecasts run over two months (Dec. 2021 – Jan. 2022). The shaded regions show the 95% confidence intervals calculated using a bootstrap method.

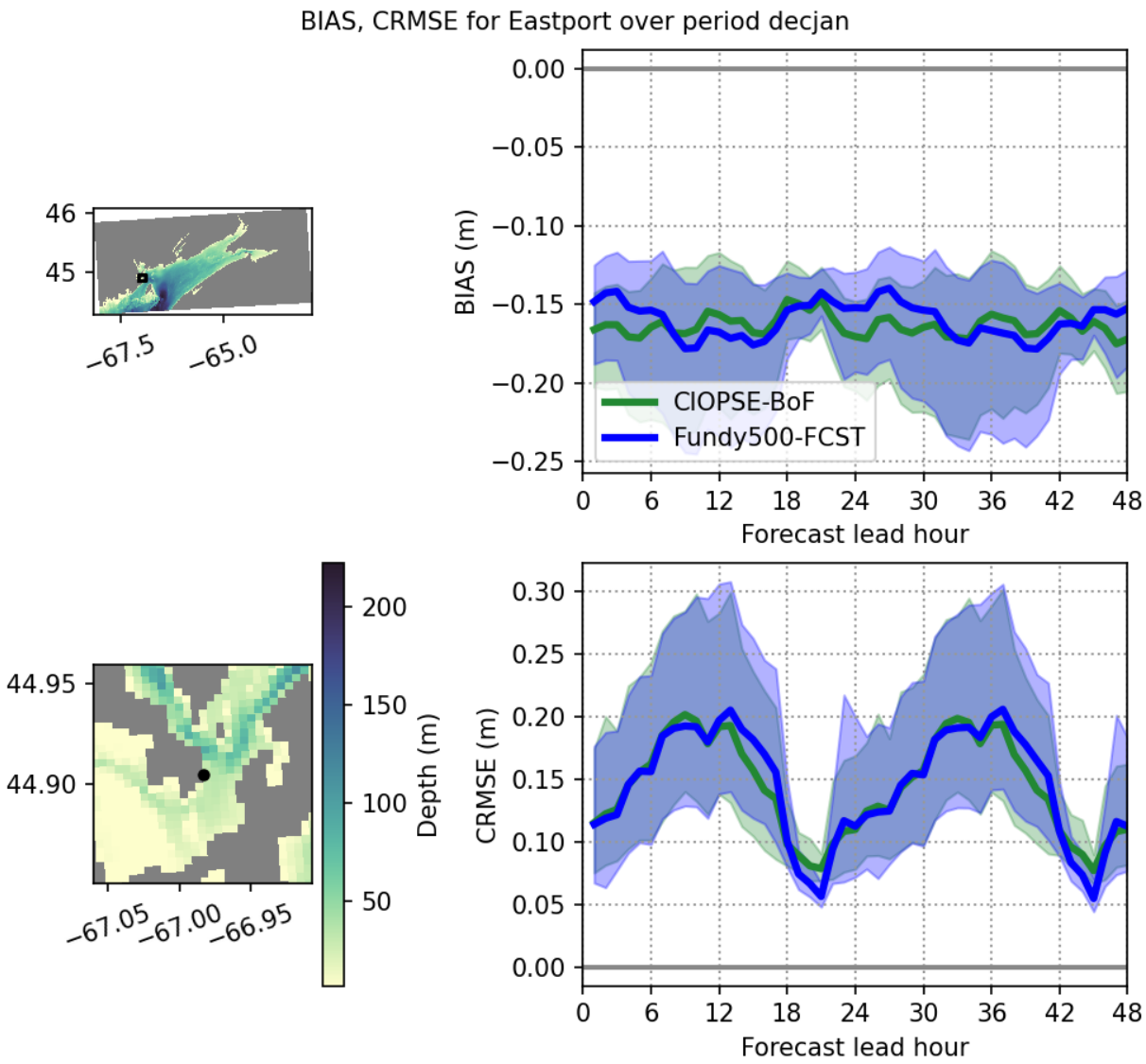

Figure 84: Non-tidal water level bias and CRMSE as a function of lead hour at Eastport for the batch of 48-hour long 00Z forecasts run over two months (Dec. 2021 – Jan. 2022). The shaded regions show the 95% confidence intervals calculated using a bootstrap method.

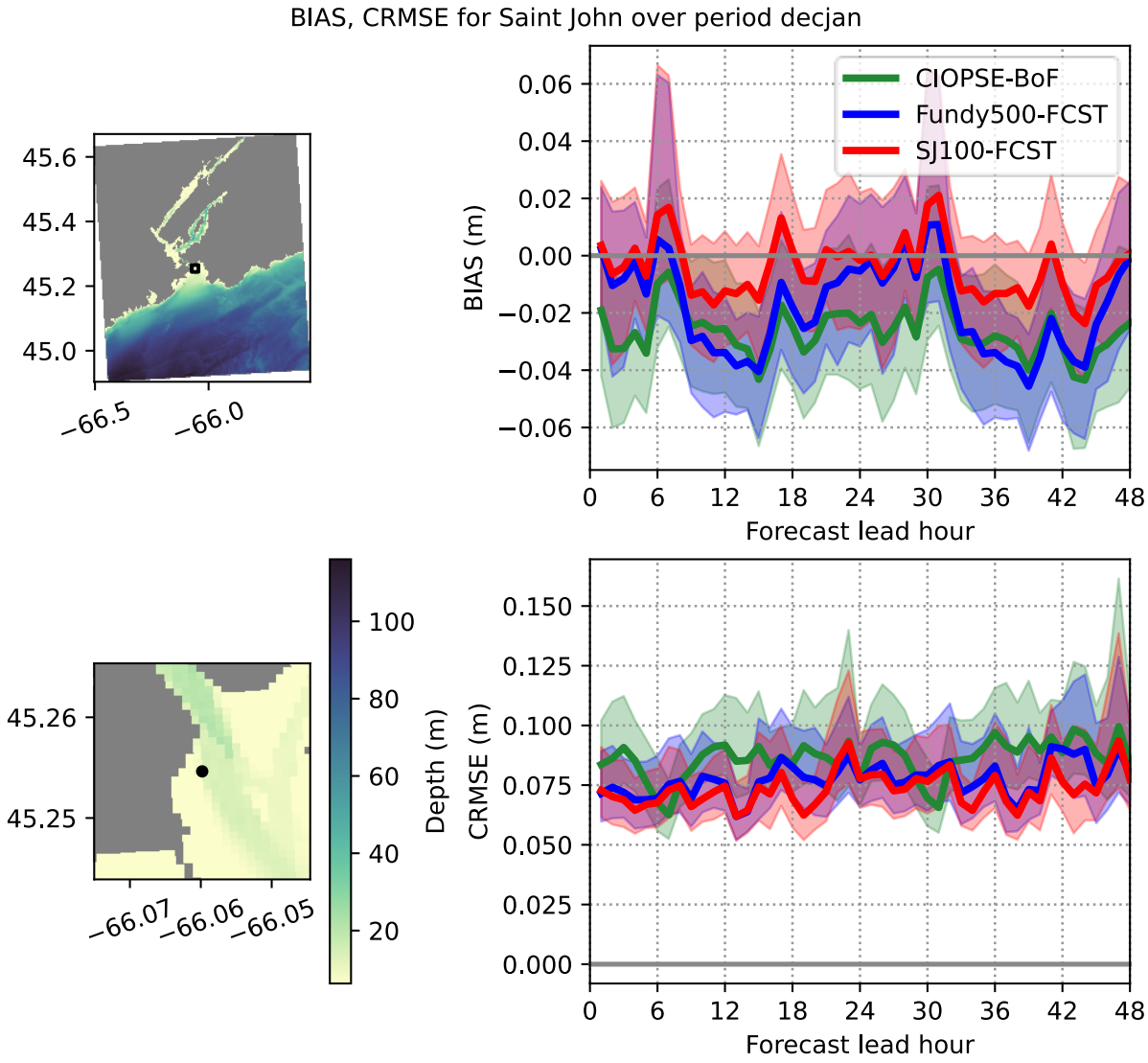

Figure 85: Non-tidal water level bias and CRMSE as a function of lead hour at Saint John for the batch of 48-hour long 00Z forecasts run over two months (Dec. 2021 – Jan. 2022). The shaded regions show the 95% confidence intervals calculated using a bootstrap method.

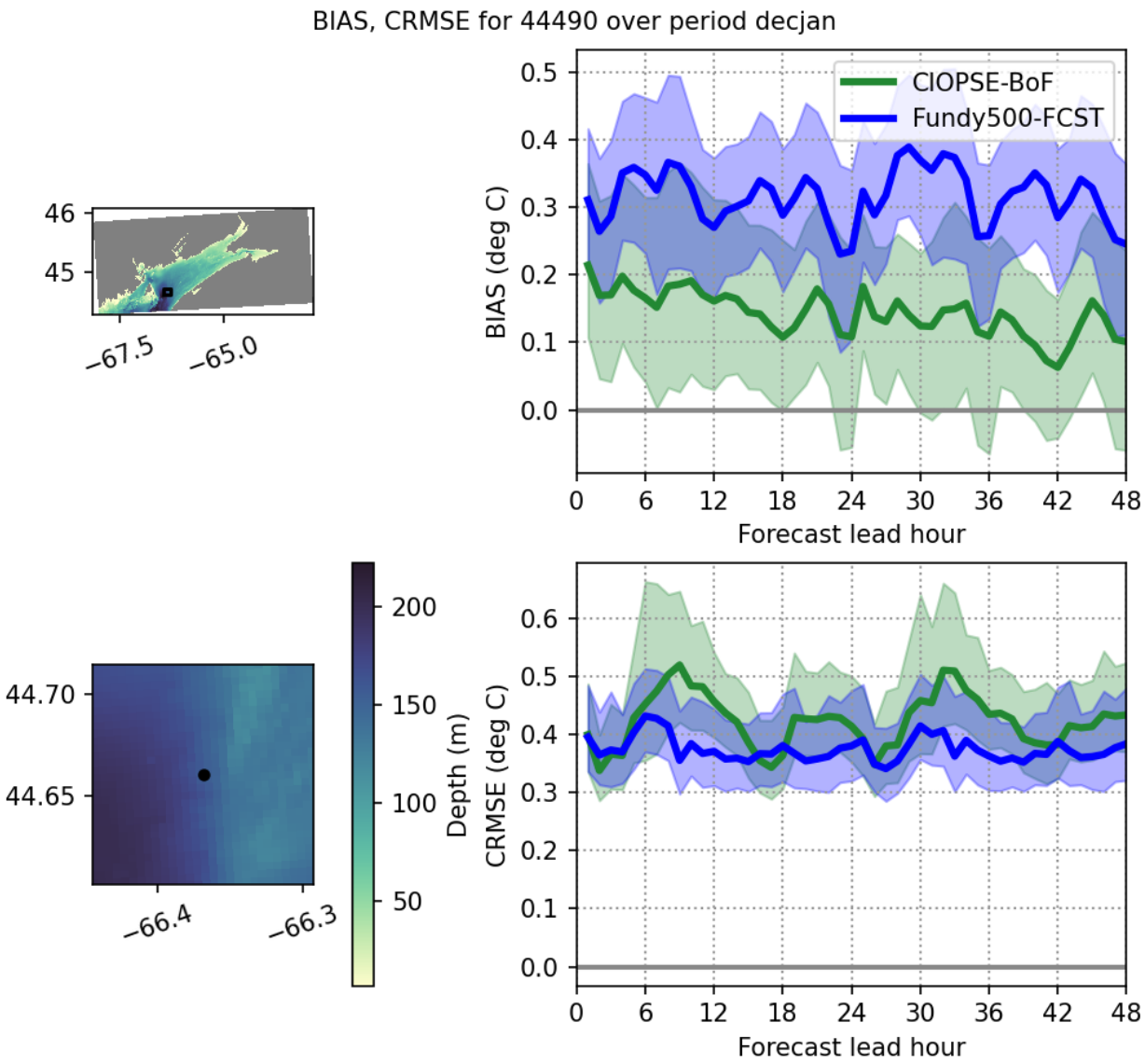

Figure 86: SST bias and CRMSE as a function of lead hour at buoy 44490 for the batch of 48-hour long 00Z forecasts run over two months (Dec. 2021 – Jan. 2022). The shaded regions show the 95% confidence intervals calculated using a bootstrap method.

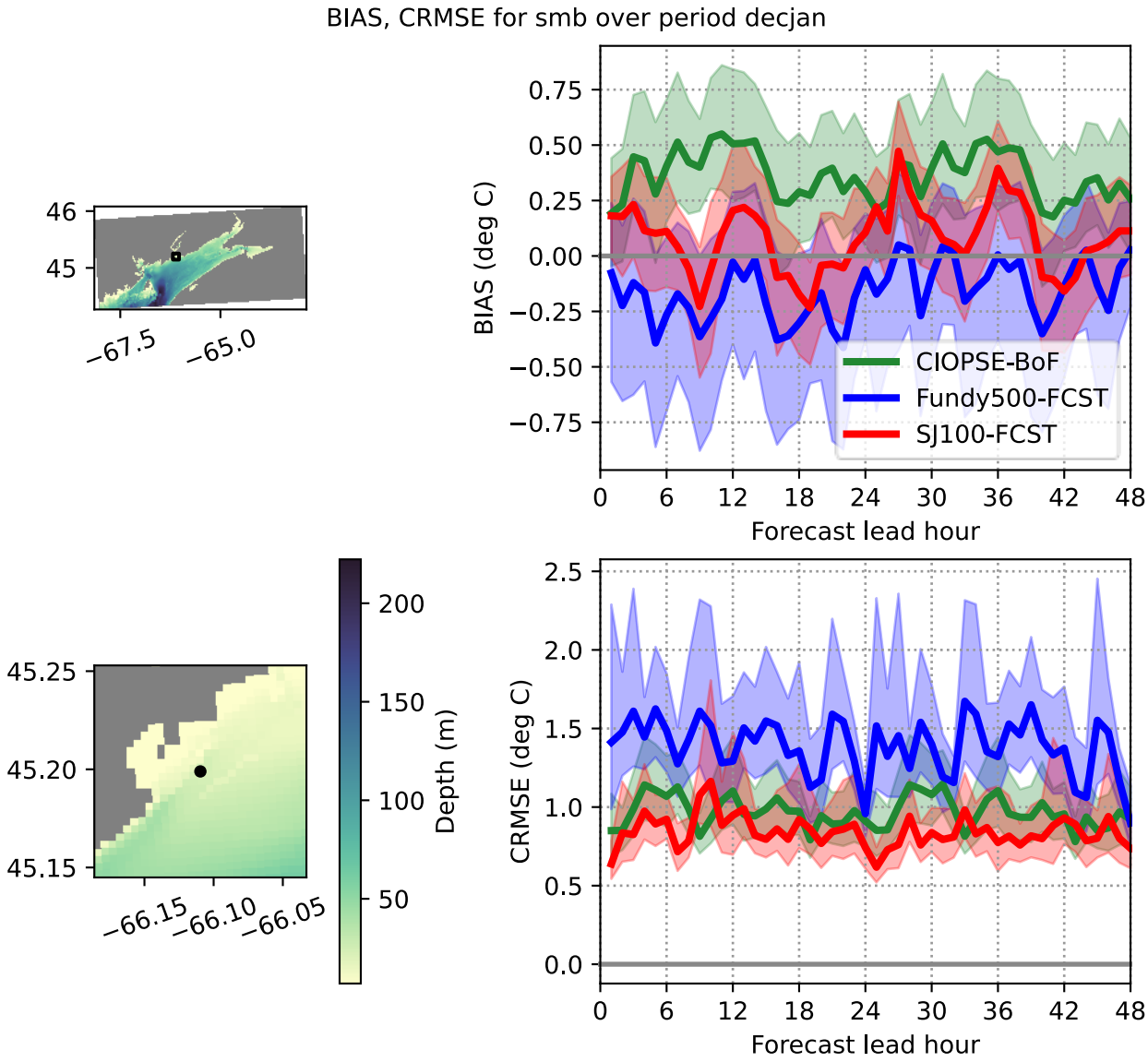

Figure 87: SST bias and CRMSE as a function of lead hour at the Smart Buoy for the batch of 48-hour long 00Z forecasts run over two months (Dec. 2021 – Jan. 2022). The shaded regions show the 95% confidence intervals calculated using a bootstrap method.

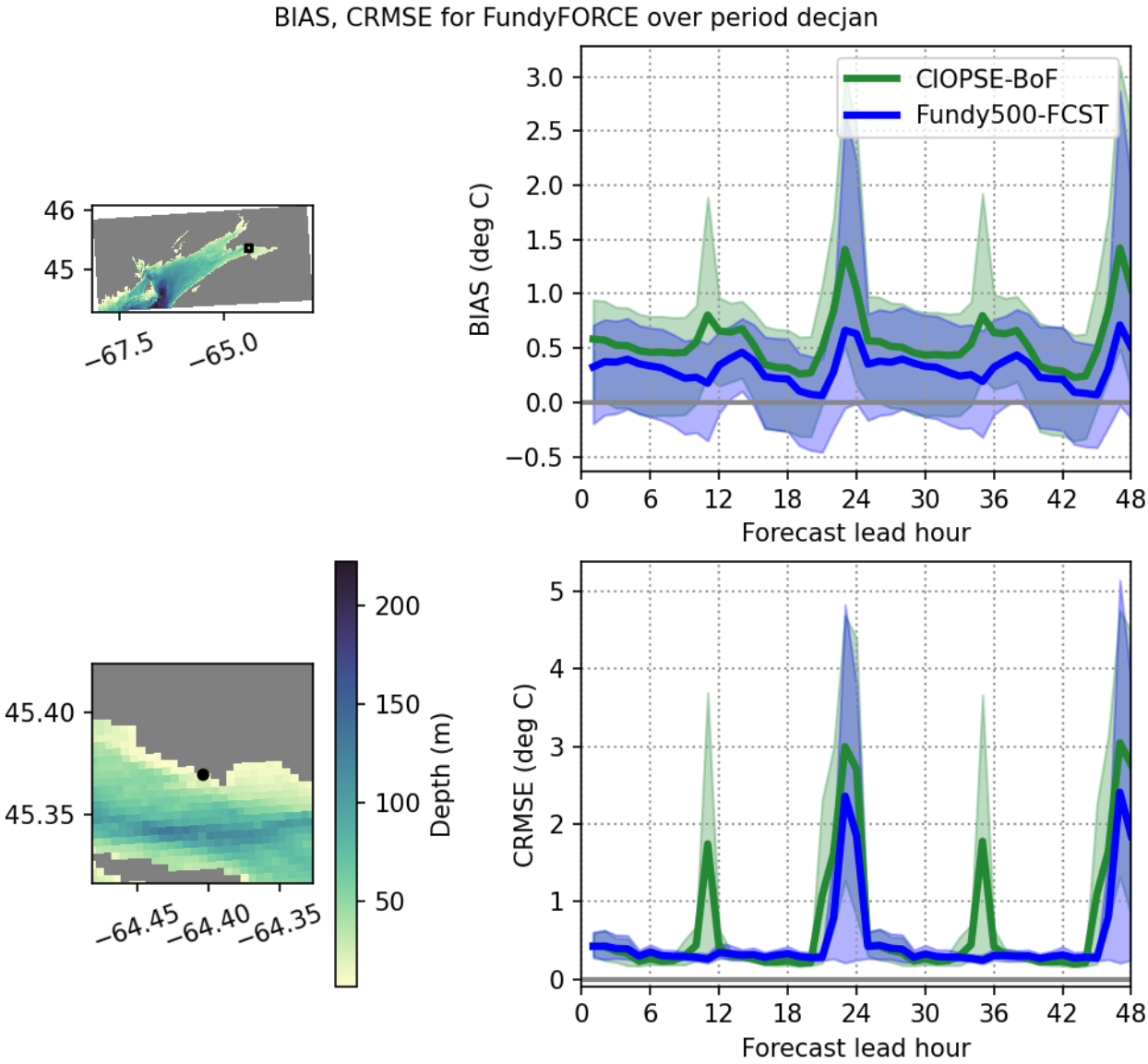

Figure 88: SST bias and CRMSE as a function of lead hour at FORCE for the batch of 48-hour long 00Z forecasts run over two months (Dec. 2021 – Jan. 2022). The shaded regions show the 95% confidence intervals calculated using a bootstrap method.