NOAA Technical Memorandum NOS 37
NOAA Technical Memorandum NWS 05
NOAA Technical Memorandum OAR 05

UFS Coastal Applications Team Report
Round 2 Summary of a Unified Forecast System Model Evaluation for Marine Navigation

Silver Spring, Maryland
April 2025

noaa National Oceanic and Atmospheric Administration

U.S. DEPARTMENT OF COMMERCE
National Ocean Service
Coast Survey Development Laboratory

Office of Coast Survey
National Ocean Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

The Office of Coast Survey (OCS) is the Nation's only official chartmaker. As the oldest United States scientific organization, dating from 1807, this office has a long history. Today it promotes safe navigation by managing the National Oceanic and Atmospheric Administration's (NOAA) nautical chart and oceanographic data collection and information programs.

There are four components of OCS:

> **The Coast Survey Development Laboratory develops new and efficient techniques to accomplish Coast Survey missions and to produce new and improved products and services for the maritime community and other coastal users.**
>
> **The Marine Chart Division acquires marine navigational data to construct and maintain nautical charts, Coast Pilots, and related marine products for the United States.**
>
> **The Hydrographic Surveys Division directs programs for ship and shore-based hydrographic survey units and conducts general hydrographic survey operations.**
>
> **The Navigational Services Division is the focal point for Coast Survey customer service activities, concentrating predominately on charting issues, fast-response hydrographic surveys, and Coast Pilot updates.**

NOAA Technical Memorandum NOS 37
NOAA Technical Memorandum NWS 05
NOAA Technical Memorandum OAR 05

UFS Coastal Applications Team Report
Round 2 summary of a Unified Forecast System Model Evaluation for Marine Navigation

Ayumi Fujisaki-Manome[1,2] (Team Co-Lead), Greg Seroka[3] (Team Co-Lead), John Kelley[3] (Team Co-Lead), Shachak Pe'eri[4] (NOS Co-Lead), Joe Sienkiewicz[5] (NWS Co-Lead), Jesse Feyen[2] (OAR Co-Lead), Olivia Doty[1], Kayo Ide[6], Brendan Gramp[6], Fred Ogden[7], Tracy Fanara[8], Edward Myers[3], Saeed Moghimi[3], Timothy Cockerill[9], Wei Wu[3], Eric Anderson[10], Kaitlin Huelse[10], Saeed Memari[10], Cristina Forbes[11], Yonggang Liu[12], Sebin John[12], Emanuele Di Lorenzo[13], Kyungmin Park[14], Spenser Wipperfurth[15], Natalia Sannikova[16], Vasily Titov[17], Yong Wei[18], Cigdem Akan[19], Soroosh Mani[3,20], Carolyn Lindley[21], Ilya Rivin[21]

1. University of Michigan/CIGLR
2. NOAA/OAR/GLERL
3. NOAA/NOS/OCS
4. NOAA/NOS/NGS
5. NOAA/NWS/NCEP/OPC
6. University of Maryland, College Park
7. NOAA/NWS/OWP
8. NOAA/NOS/IOOS
9. UT Austin/TACC
10. Colorado School of Mines
11. United States Coast Guard Search and Rescue, USCG-SAR
12. University of South Florida
13. Brown University
14. DOE/Pacific Northwest National Laboratory
15. Georgia Tech University
16. Cooperative Institute for Marine and Atmospheric Research, NOAA/OAR/PMEL
17. NOAA/OAR/PMEL
18. University of Washington
19. University of North Florida
20. SFI
21. NOAA/NOS/CO-OPS

April 2025

noaa **National Oceanic and Atmospheric Administration.**

U. S. DEPARTMENT OF COMMERCE
Howard Lutnick,
Secretary

National Oceanic and Atmospheric Administration
Laura Grimm,
Chief of Staff
Performing the Duties of Under Secretary of Commerce for Oceans and Atmosphere and NOAA Administrator

National Ocean Service
Nicole LeBoeuf,
Assistant Administrator

Office of Coast Survey
Rear Admiral Benjamin Evans
Director

Coast Survey Development Laboratory
Corey Allen
Division Chief

NOTICE

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

This report documents the second round of the model evaluation effort of the National Oceanic and Atmospheric Administration (NOAA) Unified Forecast System (UFS) Coastal Application Team (CAT) Marine Navigation Sub-Application. The effort follows the 1) the first phase of the project, which included gathering user requirements, generating an initial list of oceanographic models to evaluate, and defining skill assessment guidelines for the future model evaluation (Seroka et al. 2022), and 2) the first round of the second phase of the project– the model evaluation (Seroka et al. 2024). The project has worked toward the major goals of UFS CAT, which are 1) to evaluate potential coastal ocean models for the coastal ocean model components of the UFS and 2) to train the next generation of coastal modelers to accelerate the Research-to-Operations (R2O) process within the National Ocean Service (NOS).

Similarly to the first round of the model evaluation, two coastal ocean models were used for evaluation: 1) Finite Volume Community Ocean Model (FVCOM); and 2) Semi-implicit Cross-scale Hydroscience Integrated System Model (SCHISM). These two ocean models to date have demonstrated sufficient skill to meet NOS forecast skill requirements. Therefore, specific model results are not shown in this report; some testers' results can be found in journal publications that have resulted from this work. Instead, the report focuses on successes, challenges, and lessons learned that aid future advancement of these models in NOAA Readiness Levels as potential candidates for the UFS' operational coastal ocean model components.

In the second round, the researchers participating in the evaluation effort (i.e., testers) refined the model mesh for 3D simulations and incorporated atmospheric forcing in a 3D (multiple vertical layers) baroclinic mode whereas the first round focused on 2D simulations and tidal forcing only. The testers used atmospheric forcing from commonly used models [i.e., NOAA's High Resolution Rapid Refresh (HRRR), Global Forecast System (GFS), and/or the ECMWF Reanalysis v5 (ERA5)]. In addition, the testers used output from NOAA's Global Real-Time Ocean Forecast System (G-RTOFS) or datasets from Copernicus Marine Environment Monitoring Service (CMEMS) and Hybrid Coordinate Ocean Model (HYCOM) in order to create subtidal ocean lateral boundary conditions. For river boundary conditions, testers used river discharge observations at U.S. Geological Survey (USGS) river gauges. The model predictions were evaluated against available water levels, currents, temperature, and salinity observations during January 1 - March 31, 2022 and July 1 - September 30, 2021.

The key steps of this round two model evaluation include:

- Refine mesh for three-dimensional baroclinic simulations
- Conduct baseline simulations
- Compare model results with observations (water levels, water currents required; water temperature and salinity optional)
- Explore additional results to each tester's baseline simulation
- Conduct skill assessment based on NOAA's model evaluation guidance

Successes from the second round include continued training of the next generation of ocean modelers, testers' learning of the models and their exposure to NOAA's operational processes,

skill assessment using multiple metrics, and multiple conference presentations and publications out of the effort. Beyond improving collaboration and cooperation between modeling groups within the government and academic partners, the UFS CAT was able to promote an evaluation approach that uses multiple hydrodynamic models. By working together and building a coalition, the UFS CAT team members were able to track, support, and assess new technological advancements and algorithms pertaining to oceanographic models. Challenges include ensuring consistency and troubleshooting support among the testers while each tester has their own baseline, issues with Digital Elevation Models (DEMs), and the need for local knowledge to interpret both the inputs (e.g. DEMs) and outputs (i.e. simulations results) accurately. These successes and challenges prompted three lessons learned: 1) gap analysis (i.e. evaluating gaps in each model's performance and skill to bring the models to a "level playing field" for NOAA operations), 2) realistic expectations (i.e. testers have focused so far on setting up/learning model configurations and gaining simulation accuracy, with a lack of attention to model efficiency, requiring co-leads to adapt and emphasize efficiency in subsequent rounds), and 3) update to skill assessment software (i.e. further motivation for the ongoing effort of NOS's development of a next generation skill assessment software to have more consistency and usability, and allow for process-based skill assessment). These lessons learned will be implemented in the next rounds that focus on incorporating wave and hydrologic processes in the simulations, and testing of the UFS-Coastal infrastructure.

Overall, the work with the second round helped the UFS CAT team and NOAA's operational coastal ocean forecasting enterprise make substantial progress toward NOAA's central goals of "Building a Weather Ready Nation" and "Accelerating Growth in an Information-Based Blue Economy."

1. BACKGROUND

NOAA's Unified Forecast System (UFS) is an agency-wide effort that seeks collaboration between the US government and academia, with one of the goals to develop a standardized model evaluation process that would improve NOAA's coastal ocean forecast services for the public. The desired goal of the UFS Coastal Applications Team (CAT) is to promote an evaluation approach that is agnostic to the coastal ocean model in a coupled Earth system modeling framework. In order to evaluate progress of the effort and define the next tests, it is critical to perform a gap analysis, which will guide future developments of the different models and the coupling infrastructure (i.e., UFS-Coastal) to ensure that these models and the coupling infrastructure reach the same readiness level for an operational coupled model capability. The specific task for the UFS Coastal Applications Team was to evaluate open-source community models that could support at least one water forecast application among the UFS CAT sub-application teams (i.e., safety of navigation, risk reduction, and forecasting for flooding and inundation).

Over the past three years (2021-2024), a strong collaboration has formed between UFS CAT members from different modeling groups within the government and academic partners that use various methods and algorithms. The UFS CAT members provide a well-rounded perspective for achieving an agnostic model evaluation process. Experience of the members ranges from ocean modeling research to 24/7 operation groups within NOAA, as well as other agencies, such as the U.S. Coast Guard's Search and Rescue, Navy's Naval Research Laboratory, Pacific Northwest National Laboratory, U.S. Army Corps of Engineers Engineer Research and Development Center, and the U.S. Geological Survey's Woods Hole Coastal and Marine Science Center. In addition, the team includes leading researchers from universities, ranging from the University of Massachusetts-Dartmouth to William and Mary, University of Michigan, Brown University, University of Hawaii, Colorado School of Mines, University of Maryland College Park, Oregon State University, University of North Florida, University of South Florida, University of Washington, and Georgia Institute of Technology.

As a first step, the UFS CAT tasked three working groups (sub-application teams) to generate consensus guidelines (i.e., metrics, criteria, and competing numerical oceanographic models) for a model evaluation based on the application. The operational requirements for model selection for the safe, efficient marine navigation sub-application included (see requirements for all three sub-application teams here: Seroka et al., 2022; Van der Westhuysen et al., 2022; Huang et al., 2022):

- Stability and computational efficiency
- Accuracy
- Resolution
- Code management
- Coupling
- Community support and license type
- NOAA Readiness Levels
- Geographic coverage

In the case of the marine navigation (Safe, Efficient Navigation) sub-application team, the key user variables that were identified included: water levels, surface water currents, sea and lake ice, and water temperature and salinity. It is important to note that surface wind and waves were also considered as required user variables for marine navigation.

With the help of University of Texas, Austin's High Performance Computing (HPC) servers at Texas Advanced Computing Center (TACC), members from the UFS CAT from government and academia were able to work together in the same IT environment, assess the performance of model candidates, and share results between the teams. After providing criteria for an objective selection of an oceanographic model(s), four models were identified as potential candidates for evaluation. Out of the four candidate models identified as suitable for the UFS marine navigation community, two models were selected for a full evaluation in different configurations, based on NOAA operational configuration guidelines that include computational efficiency, resolution, geographic coverage spanning complex shorelines, and the ability to provide surface currents for navigation (Seroka et al., 2022). Three-dimensional unstructured grid hydrodynamic models are thus the most effective at meeting these and the other operational requirements:

1) Finite Volume Community Ocean (FVCOM) Model (Chen et al., 2002; 2006). FVCOM is currently used in the operational National Ocean Service (NOS) Great Lakes forecast systems, Northern Gulf of Mexico forecast system, San Francisco Bay forecast system, and is under development for some Atlantic and Pacific regional models

2) Semi-implicit Cross-scale Hydroscience Integrated System (SCHISM) Model (Zhang et al., 2016). The SCHISM model is currently used for the 3D Surge and Tide Operational Forecast System in the Atlantic Ocean (including Gulf of Mexico and Caribbean Sea, STOFS-3D-Atlantic), is also being applied in the development of the Pacific 3D STOFS (STOFS-3D-Pacific), and is under development for a model system for the Southeast U.S. and Puerto Rico.

The two models were evaluated independent from each other using multiple configurations (rounds) with a unique objective for each round. With the recognition that UFS CAT testers come from different universities with varying experience, the first round allowed the testers to learn how to develop model meshes, how to prepare, make, and evaluate tidal simulations. The main objective in this round was team building. **The models were evaluated in round one in average water density (2D) mode and were not coupled to other models**, where the water level (tidal) components were evaluated. This round is completed with each team member able to properly install the circulation models on the TACC environment (with the help of the model developers), make and evaluate tidal simulations to ensure quality 3D model simulations in subsequent rounds, and get familiar with the model.

Recently, the testers completed the second round of evaluation, where the teams evaluated the models in a layered water density (3D) mode, forced with atmospheric models (GFS/HRRR). In addition to water levels, the testers also evaluated surface currents (top 4.5 m water layer), water temperature and salinity. This report provides a summary of the second round setup and results, including an update on the model meshes, calibration procedures, and an expansion of the skill assessment to include additional observational data (water levels, water currents, water temperature, salinity). Key steps of this round of model evaluation include:

- Refine mesh (unstructured grid) for three-dimensional baroclinic simulations
- Conduct baseline simulations
- Compare model results with observations (water levels, water currents required; water temperature and salinity optional)
- Explore additional results to each tester's baseline simulation
- Conduct skill assessment based on NOAA's model evaluation guidance

While this second round work does not include coupling between the hydrodynamic model and other environmental models, the outcomes of this round create the infrastructure for the next rounds of evaluation, where coupling using the UFS-Coastal infrastructure will be tested. In addition to testing the UFS-Coastal infrastructure, the future rounds will also include incorporating wave processes (e.g. WaveWatch III) as well as hydrologic processes (e.g. National Water Model).

The development and implementation steps of the UFS-Coastal framework (Figure 1) include:

- Ongoing development of the UFS-Coastal code base forked from ufs-weather-model by including coastal ocean modeling components.
- Ongoing development of automated regression testing cases for coastal ocean model components. This also includes the UFS CAT members testing the automated regression tests for their applications, and providing feedback to the UFS-Coastal developers.
- Help inform future "Case Studies" at the ufs-coastal-app level based on more realistic cases with higher resolution meshes and more forcing variables developed by the UFS CAT testers
- Utilizing the ocean modeling github organization as the community platform to jointly develop UFS-Coastal and related applications (e.g. marine navigation, risk reduction and total water level; see: https://github.com/oceanmodeling/ufs-weather-model).
- The UFS CAT members ask questions and receive feedback and support on Github Discussions and Issues. Virtual office hours with the UFS-Coastal Team also provides a venue for the members to obtain support from the UFS-Coastal Team.
- Establishing a Continuous Integration and Continuous Delivery (CI/CD) pipeline to employ software development best practices

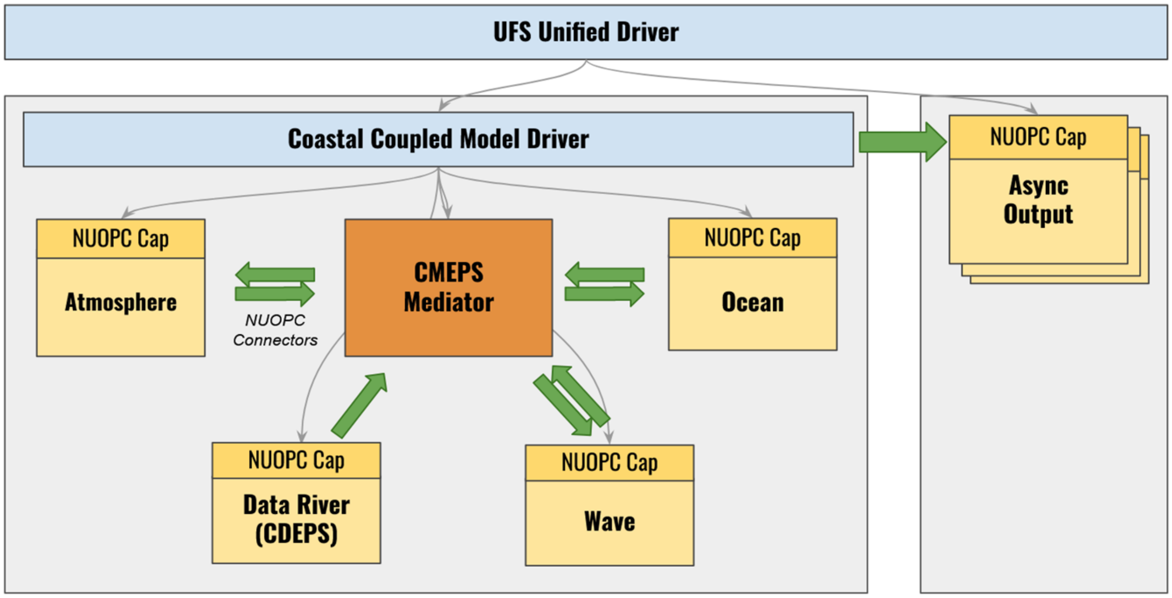


Figure 1. Envisioned UFS-Coastal coupling code infrastructure diagram 1

2. MODEL EVALUATION REQUIREMENTS

Following UFS CAT Marine Navigation Water Quantity white paper (NOAA, 2022), the model evaluation requirements (Table 1) are based on and collected from users, such as commercial and recreational mariners, port authorities, National Weather Service (NWS) and private forecasters, marine educators/researchers, national and state level search and rescue, manufacturers of marine navigational systems, and offshore wind energy operators. The accuracies of these requirements (Table 2) align with the International Hydrographic Organization (IHO) that has also been collecting user requirements in order to create product standards (IHO S-1xx) to be used as part of a carriage suite on certain vessels that can be displayed on an Electronic Chart Display Information System (ECDIS).

Table 1. Forecast configuration requirements of circulation model evaluation for Safe and Efficient Navigation (NOAA, 2022)

Key user variables:	**Specifications**
Forecast frequency	Every 6 hours
Forecast turnaround time	< 1 hr before forecast cycle deadline (NWS), and before start of the next model forecast cycle (NOS)
Output temporal resolution	At least hourly, optimally 6 minutes
Forecast range	5 to 7 days; 14 days for planning (monthly/seasonal for lake/sea ice) were considered but not detailed here
Reliability	99-99.9%
Areas of interest	Coastal ocean, Great Lakes, including ports, harbors, bays, and connecting channels and rivers, and islands/atolls in the Pacific (e.g. Hawaiian Islands and Guam)
Depth of currents	Navigation - 4.5 m below surface Search and Rescue - 0-1 m below surface
Spatial reference system	Vertical - Chart datum (e.g. MLLW and LWD for Great Lakes) Horizontal - WGS-84 or ITRF
Horizontal resolution	Rivers - 10 m in rivers, Shipping channels - 10s of m in shipping channels, Sea ice conditions - 30 m for sea ice,

	Within inlets, bays and lakes - 50 m-1km, Around small islands - <=2 km, Open ocean conditions - 5 km (1 km for surface currents in EEZ) It is also important to represent coastal shoreline structures, such as levees, piers, offshore wind farms

Table 2. Accuracy requirements of key variables used in the circulation model evaluation for Safe and Efficient Navigation (NOAA, 2022)

Key user variables:	Specifications and Accuracy (acceptable Root Mean Square Error):
Water level accuracy	Under Keel Clearance (UKC) - 15 cm (0.5 ft) Time of high water and time of low water - 0.5 hr
Surface current accuracy	Speed - 26 cm/sec (0.5 kt); at time of max flood or ebb 30 min; for slack water times, 15 min (Note: For USCG-SAR: 10 cm/sec, 0.2 kt) Direction: 22.5 degrees provided current speed is not less than 26 cm/s (0.5 kt) (Note: For USCG-SAR: 10 degrees)
Sea and lake ice accuracy	Depth/thickness - 10 cm Concentration - 10% Extent - 10% Motion - 0.25 km/day / 10 degrees
Water density accuracy	Salinity - 3.5 psu for salinity Water temperature - 7.7C (Note: Desired accuracy is to forecast a ship's draft within 7.5 cm of its actual draft).
Product formats:	S-100/HDF5, GRIB2, Web mapping services, GIS compatible files, NetCDF, SHEF; documentation describing files/variables

3. DETAILS OF MODEL EVALUATION AND PROCESSES

3.1. Selection criteria

Both coastal ocean models, FVCOM and SCHISM, utilize free surface 3D primitive equations on unstructured grid frameworks. They both have the capability to expand to a fully coupled current-ice-wave-sediment-ecosystem model with parallelization implemented. This type of 3D circulation model allows for accurate simulations of 3D baroclinic circulation, spanning from small waterways and islands to ocean-wide patterns, all within a single grid.

As described in UFS CAT Marine Navigation working group white paper (Seroka et al., 2022), the operational guidelines for coastal models for marine navigation applications include:

Stability and computational efficiency - Computation and delivery of model products need to be fast enough to provide forecasters and users actionable information in a timely manner. Numerical speed and stability that can operate on HPC infrastructures are available at NOAA's NCEP Central Operations (NCO). Overall, the model robustness is a key requirement for operational use. Operational grade models are expected to run reliably, and provide guidance for each cycle without fail.

Accuracy - The accuracy should be defined based on physical and not numerical adjustments. Observations that are accepted by NOAA should be used for skill assessment. This also requires that delivered products are free from erroneous behavior such as spurious water level or wave height peaks.

Resolution - An operator should be able to use a new bathymetry grid and generate a mesh at the horizontal resolution defined in the configuration forecast. An additional operational requirement for resolution is the sensitivity to bottom slope or required topobathy smoothing.

Code management - Even if NOAA is running the models operationally and the operational code is differentiated from developmental codes, it is very important that the code is managed and developed by a scientific community. It is important to make the models available to the scientific community (i.e., developers) through github or similar git platforms, and for NOAA to integrate updates into the UFS code base.

Coupling - The models need to be able to couple to ocean, wave, inland hydrology, atmosphere, and sea ice models. It is important that the circulation model is using the National Unified Operational Prediction Capability (NUOPC) cap in order to inform and receive information from other models (such as atmospheric, wind, ice, and hydrology) and that the circulation model is coupled using the UFS driver.

Community support and license type - All coupled model components are required to be community models that are open source (License CC0 or equivalent) and supported by an active user community. If a License CC0 is not available, the models need to have an open source/open access to government and not-for-profit groups for allowing changes to be made and distributed, at a minimum.

NOAA Readiness Levels - NOAA has adopted a systematic project metric/measurement system that supports assessments of the maturity of R&D projects from research to operation, application, commercial product or service, or other use and allows the consistent comparison of maturity between different types of R&D projects. Model candidates are considered to be part of a UFS application based on a portfolio of research contributions. UFS models and their components follow the NOAA Readiness Levels guidance to determine operational readiness and implementation (NOAA, 2025a,b).

Geographic coverage - The models should be able to operate successfully in coastal environments that include all of the United States top 50 ports (BTS, 2025). The geographic settings of some of these ports are complex (up a major river with a tidal signal, complex shoreline, varying tidal ranges).

3.2. Team Development

With the understanding that the UFS CAT includes different teams that undergo team building at different paces, the team embraced the developmental stage processes (Tuckman, 1965) and, as a group, assisted team members transitioning between the different stages. The transitions included moving from being in a comfort zone of non-threatening topics and avoiding the risk of conflict (“Forming”), to "testing and proving" mentality (“Storming”) to a problem-solving mentality (“Norming”). Eventually, the team evolved into its current capacity, range, and depth of personal relations, and expanded to true interdependence (“Performing”). In this last stage mentioned, people within the team can work independently, in subgroups, or as a total unit with equal competencies.

While the CAT members were transitioning as a group from the “forming” phase into “storming” and “norming” phases in Round 1, the CAT members are now “performing” in Round 2. The co-leads continued to work with the model developers, Dr. Chen from University of Massachusetts Dartmouth and Dr. Zhang from William & Mary (Virginia Institute of Marine Sciences), and University of Texas, Austin (UT Austin) TACC to maintain a HPC environment for each of the models. The goal of Round 2 was to continue model evaluations using the circulation models on the UT Austin/TACC environment (with the help of the model developers) and facilitate the testers’ learning environment for the models. The testers included several faculty members and students from the following universities:

- Brown University
- University of Maryland, College Park
- University of Hawaii
- Colorado School of Mines
- University of South Florida
- University of Michigan
- University of North Florida

In addition, Federal Employees and contractors from several groups at NOAA participated in the model evaluation. At first, access was provided to the model developers and testers to compile and test the model candidate on the TACC infrastructure using standard test cases. After working with the model developers to optimize the compilation given standard compilers and libraries on UT Austin//TACC to a reliable and performance state, the testers were onboarded and standard test runs were performed for a given region and a given specified computational time / resources to allow fair quantitative model evaluation. The total requested node hours for this start up project is 5000 Service Units (SUs).

Followed by this start up project, a larger allocation was requested and awarded, with a total of 272800 SUs, for more computationally expensive 3-dimensional simulations in Round 2, which were forced by atmospheric data. In Round 2, the key steps of the model evaluation include:

- Refine mesh for three-dimensional baroclinic simulations
- Conduct baseline simulations
- Compare model results with observations (water levels, water currents required; water temperature and salinity optional)
- Explore additional results to each tester's baseline simulation
- Conduct skill assessment based on NOAA's model evaluation guidance

3.3. Mesh Generation

The study site selected for model evaluation is New York Harbor that is considered one of the top 5 ports in trade (cargo weight) within the US. Entrance to the port is from the Atlantic Ocean into Raritan Bay, located between New York and New Jersey. Within the Bay, there are two main channels leading into the harbor (Ambrose Channel and the Sandy Hook Channel/Raritan Bay reaches), and several inland rivers connected to the bay, including Hudson River, East River, Arthur Kill, Raritan River, and Hackensack River. Emphasis in the model processing was placed on mesh-grid development by each team, with the use of a provided ('blessed') mesh allowed for those who have computation or time restraints.

A compiled high-resolution bathymetry dataset of the study site (from land up to the continental shelf) was provided for mesh generation to all teams (Figure 2). Data sources used to generate the bathymetry dataset included: NOAA Office of Coast Survey's National Bathymetric Source (NBS) (Wiley and Rice, 2020) and NCEI's Continuously Updated Digital Elevation Models (Amante et al., 2023) with a grid resolution around the study area (New York Harbor, Long Island and its surrounding rivers) at 4 m, and coastal areas around Rhode Island, Massachusetts and New Jersey at 8 m. To supplement the high-resolution bathymetry, offshore areas on the continental shelf were provided at 16 m and 32 m, and the continental shelf and slope were provided at 100 m using the NOAA hydrographic survey grids and the US Extended Continental Shelf program (Gardner et al., 2005). All the datasets were vertically referenced to the orthometric height of North American Vertical Datum of 1988 (NAVD88) and horizontally to World Geodetic Survey 1984 (WGS-84). However, due to the data sources and the volume of data, different grid resolutions were provided based on the geographic coverage.

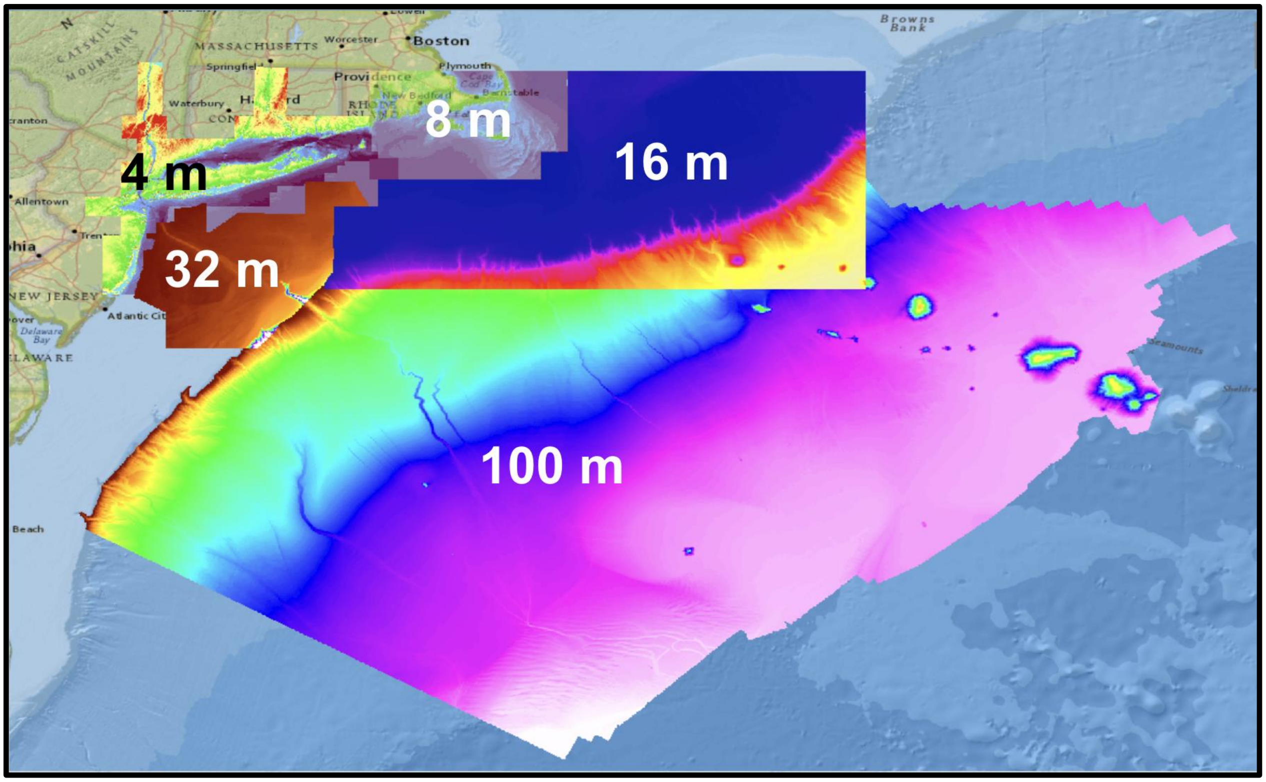


Figure 2. Available bathymetry datasets at the study site (NY Harbor) for model evaluation. The numbers over shading represent horizontal resolution of the digital elevation model (DEM).

According to the model evaluation requirements, the mesh derived from the bathymetry should resolve regions as small as 10 m horizontally. Many teams found it difficult to generate a mesh at that resolution while meeting the computational expense requirements (i.e., each three-month simulation should be within ~12 hours for tidal runs and atmospheric forcing and ~18+ hours for coupled runs with wave), Therefore, instead the teams chose to have their finest resolution between 20 meters and 80 meters. This difference in resolution between teams is a key factor in analyzing the accuracy of results, along with ensuring the waters surrounding verification stations are resolved appropriately as well. In Round 2, many testers updated their mesh for refinement and adapted from 2D to 3D configurations. The meshes can be updated and refined in future rounds as well. By running the models using the meshes generated from the bathymetry, the testers identified a couple of bathymetry data issues over the course of the model evaluation:

1) Data gaps along the Arthur Kill area (Figure 3a). This issue might have introduced shallow values (“0 m”) in the dataset’s reference system. By re-evaluating the data and working closely with the NOAA NBS, a complete dataset that was correctly vertically-referenced was used along the Arthur Kill area (Figure 3b).

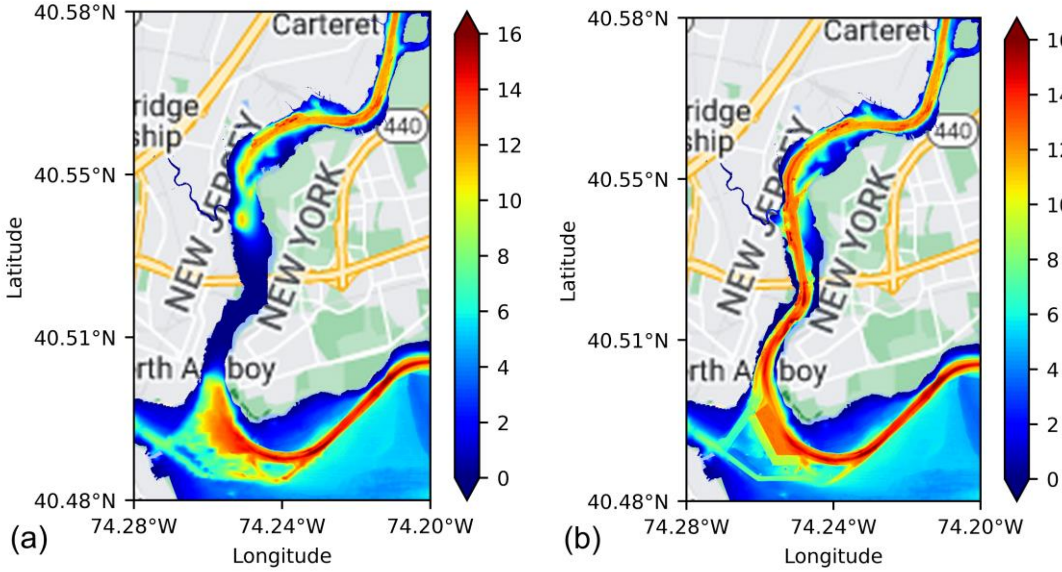
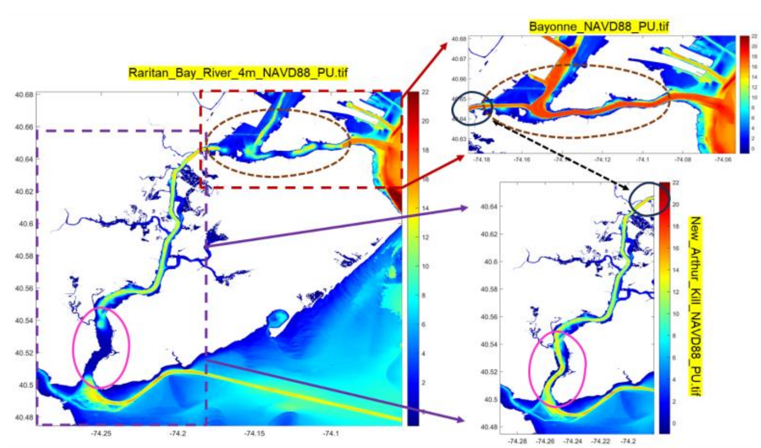


Figure 3 a) Disconnected River over the Arthur Kill, b) Fixed bathymetry data.

2) Shoal-biased areas along the Bayonne Channel and Arthur Kill (Figure 4a). These shoals might have been introduced by using a wrong reference system in one of the legacy survey datasets. These issues were remedied in the latest version of the NOAA NBS. By working closely with the NOAA NBS team, an updated vertically-referenced survey dataset was incorporated into the bathymetry dataset and provided to the UFS CAT team (Figure 4b).



(a)

(b)

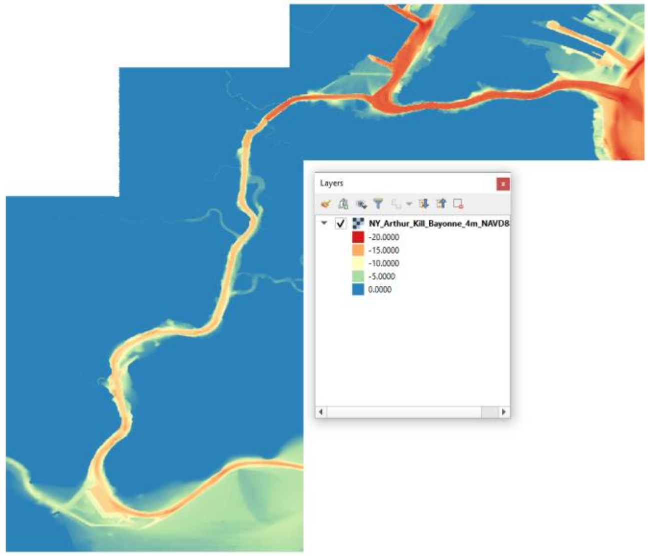


Figure 4. a) inconsistent bathymetry across the Bayonne Channel tile and the Arthur Kill tile. b) Updated bathymetry data that are consistent across the tiles.

These updates led to improvements to bathymetry representation over the areas for model simulations. The process of identifying, reporting, and fixing bathymetry data issues was a critical part of iterative mesh development, underscoring the importance of identifying Digital Elevation Model (DEM) issues early on in the mesh development (see Section 4.2 Challenges) as it impacts the model performance across all three rounds.

The tools for creating mesh from bathymetry included SMS (Surface-water Modeling System) (Militello and Zundel, 1999; Aquaveo, 2021), OCSMesh (Mani, 2021), and OceanMesh2D (Roberts, 2019). Ideally, the testers would generate their own mesh, or at least go through that exercise. With the understanding that some may struggle creating and modifying an “ideal” mesh, a mesh from other tester(s) was also provided to the testers as a backup; one mesh for SCHISM (Figure 5) and one mesh for FVCOM (Figure 6). These back-up meshes were mainly for the testers who were limited in time and resources, and for testers who developed a mesh that did not work well in the subsequent evaluation processes.

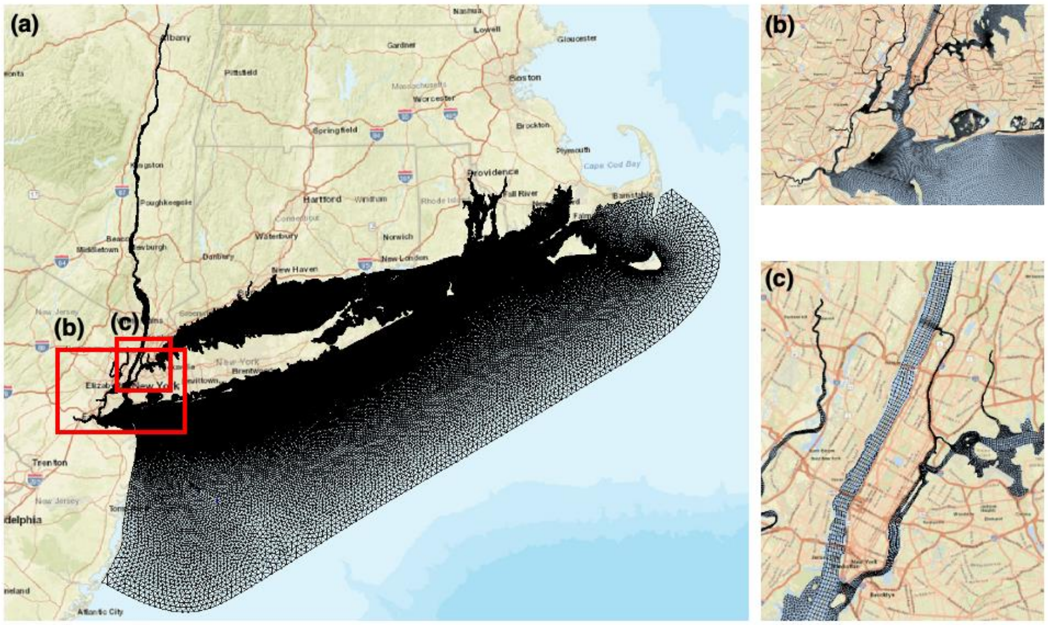


Figure 5. a) SCHISM mesh coverage (“example mesh”) and subsets containing key navigation channels: b) NY/NJ Harbor and (c) Manhattan.

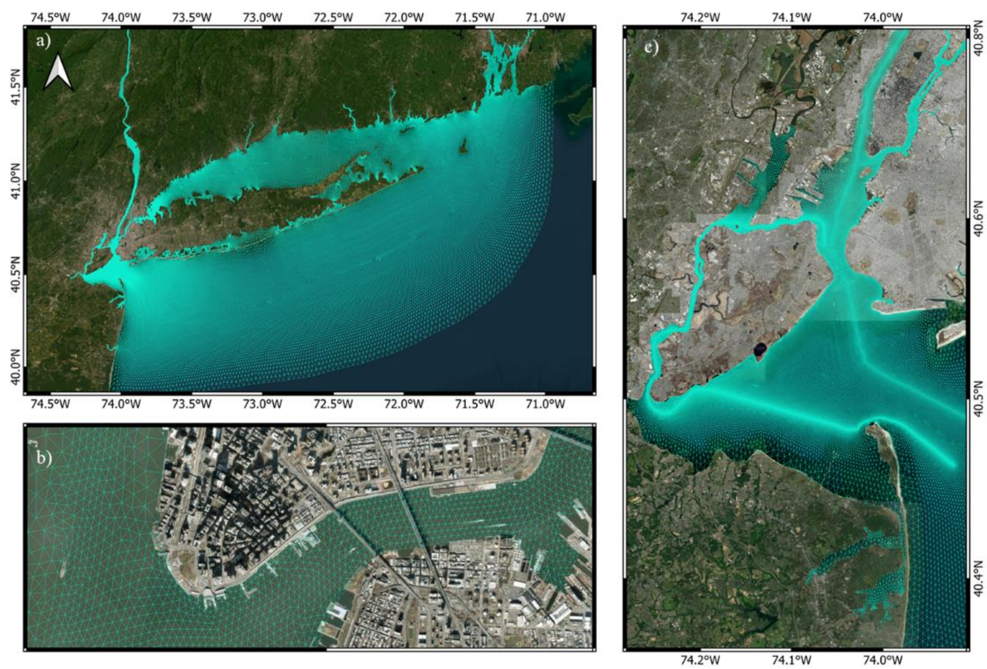


Figure 6. a) FVCOM mesh coverage (“example mesh”) and subsets containing key navigation channels: b) the area around Manhattan Island, and c) Staten Island and Brooklyn area.

3.4. Conduct baseline and other simulations

The source codes of both coastal ocean models were shared using Github repository as follows:

- **FVCOM** version 4.4.2 was used in the model configuration for the FVCOM and was shared with all testers through the GitHub repository (https://github.com/FVCOM-GitHub).
- **SCHISM** version 5.10.1 was used in the model configuration for the SCHISM and was shared with all testers through the GitHub repository (https://github.com/schism-dev/schism/). Dr. Zhang (Virginia Institute of Marine Science, VIMS) also provided the team with an updated SCHISM manual (https://schism-dev.github.io/schism/v5.10/index.html).

While Round 1 model evaluation effort focused on evaluating only the tides and 2D simulations (Seroka et al., 2024), Round 2 focused on 3D simulations forced by atmospheric data. In both rounds tidal forcing was also evaluated. Each tester conducted their own "baseline" configured simulation, which acted as a starting point for evaluating sensitivity to different forcing data, model setups, air-sea flux formulation, etc. Testers were encouraged to follow guidelines as much as possible, to allow for sharing knowledge and lessons learned among the testers.

Tidal Forcing - Most FVCOM testers used TPXO9 (Egbert and Erofeeva, 2002) or FES2014 (Lyard, 2021) as the boundary condition, where one team compared between TPXO9, FES2014, and NECOFS (Beardsley, 2014). The SCHISM tester teams used FES2014 as the boundary conditions in their model runs. One SCHISM tester also evaluated the different versions of TPXO9 (version 1 and version 5) and FES2014 (Park et al., in review). Because there was no clear cut best tidal forcing dataset to use, the co-leads allowed the flexibility for testing various options and determining the sensitivity amongst them. However, all testers were required to use eight constituents for tidal forcing:

- M2 - principal lunar semidiurnal constituent,
- S2 - principal solar semidiurnal constituent,
- N2 - larger lunar elliptic semidiurnal constituent,
- K1 - lunar diurnal constituent,
- K2 - principal lunar semidiurnal constituent,
- O1 - lunar diurnal constituent,
- P1 - solar diurnal, and
- Q1- larger lunar elliptic diurnal

Atmospheric forcing - The hourly surface meteorological outputs from NOAA's High Resolution Rapid Refresh (HRRR, Dowell et al., 2024) were used for atmospheric forcing. The provided data were from forecast hour 1 ("f01", recommended) and forecast hour 2 ("f02"). Testers needed to combine data at forecast hour 1 (or 2) in time to create temporally-continuous atmospheric forcing data for each of the two hindcast periods, including the spin-up period. In cases where a computational domain goes outside the HRRR domain, surface meteorological outputs from NOAA's Global Forecast System (GFS) were also provided. Some testers also used the

European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5 (ERA5) dataset for filling the gaps in HRRR/GFS or for testing an alternative forcing.

River forcing- Discharge datasets at five stations from the U.S. Geological Survey (USGS) were provided. They are for Hackensack River at New Milford NJ (01378500), Hudson River at Green Island NY (01358000), Passaic River at Dundee Dam at Clifton NJ (01389890), Rahway River at Rahway NJ (01395000), and Raritan River below Calco Dam at Bound Brook NJ (01403060). In addition, outputs from the National Water Model were also provided as an alternative river forcing data source.

Subtidal ocean boundary conditions- All testers were provided predictions from NOAA's Global operational Real-Time Ocean Forecast System (G-RTOFS) for the subtidal ocean boundary conditions. Some testers explored other datasets, such as data from the Copernicus Marine Environment Monitoring Service (CMEMS) and the global dataset from the Hybrid Coordinate Ocean Model (HYCOM). Forcing datasets used by tester groups are summarized in Table 3.

Table 3. Information on the data sources used by each tester. Bold river names indicate that the river forcing data was not provided to testers by co-leads.

	PNNL (SCHISM)	**PMEL (SCHISM)**	**Mines (SCHISM)**	**Mines (FVCOM)**	**UMD (FVCOM)**	**USF (FVCOM)**
Atmospheric Data Source	HRRR	HRRR	ERA5	HRRR	HRRR	HRRR
Tidal Data Source	FES2014	TPXO	FES2014	FES2014	FES2014	NECOFS
Initial Condition (Hot start)	CMEMS	HYCOM	CMEMS	GRTOFS	NECOFS	HYCOM
River Data Source	USGS	USGS	USGS	USGS	USGS	USGS
Number of Rivers	5	2	5	13	9	4
Names of Rivers	Hudson, Hackensack, Passaic, Rahway, Raritan	Hudson, Raritan	Hudson, Hackensack, Passaic, Rahway, Raritan	**Navesink**, Raritan, **Newark Bay**, Hudson, **Connecticut, Thames, Mystic, Pawcatuck, Taunton, Lee, Cole, Palmer, Providence**	Hudson, Hackensack, Passaic, Rahway, Raritan, **Manasquan, Swimming, Connetquot, Peconic**	Hudson, Passaic, Raritan, **Connecticut**

Overall, the testers have the ability to configure the settings for their convenience, but with a few requirements:

1. Hindcast Period: Three months, not including spin-up period.
    a. Jan 1 - Mar 31, 2022: includes Nor'easter events (Jan 14-17; Jan 28-29)
    b. Jul 1 - Sep 30, 2021: includes hurricane events (Elsa, Henri, Ida)

2. Run time: Each three-month simulation should be within ~12 hours (tidal runs and atmospheric forcing) and ~18+ hours (coupled with wave)

3. Mesh generation: Use provided bathymetry, geographic polygon for mesh generation. Resolve small riverine areas at ~10 m horizontal resolution.

The teams also were provided with flexibility to choose the number of cores to run their model that resulted in varying run times based on the nodal count and connectivity of their mesh (Table 4). All teams were provided access to the TACC Frontera system for testing. In Round 2, the number of nodes for each team's mesh ranged from 21,623 to 124,465.

Table 4. Configurations for each tester's computational runs and mesh refinement

	PNNL (SCHISM)	**PMEL (SCHISM)**	**Mines (SCHISM)**	**Mines (FVCOM)**	**UMD (FVCOM)**	**USF (FVCOM)**
Computational nodes	10	10	5	5	20	12
Computational cores	560	560	280	280	500	500
Runtime (hrs)	4	6.5	5	48	4	60
Seconds per core	25.7	41.8	64.3	617.1	28.8	432
Cores per hour (speed)	140	86.2	56	5.8	125	8.3
# of mesh nodes	95,566	63,105	70,161	88,842	21,623	124,465
# of mesh elements	165,132	112,542	129,921	160,515	34,876	235,931
# of vertical layers	up to 22	up to 33	31	40	20	21
(nodes*layers)/runtime in hours	525,613	320,379	434,998	74,035	108,115	43,563
Highest Resolution (inland river) (m)	5	50	67	80	67	20

3.5. Skill assessment using available observations

The comparison of the model results utilized available *In Situ* observations within the study area, including those from water level gauge stations, buoys, and Acoustic Doppler Current Profilers (ADCPs). Additional *In Situ* datasets included observations from gliders, high-frequency radar, and eXpendable BathyThermographs (XBTs). All *In Situ* datasets were available to the testers on TACC. The testers used the evaluation metrics documented in the white paper from the Navigation working group (Seroka et al., 2022). The specific skill assessment criteria for Round 2 include mandatory evaluations and additional options, based on available time:

Mandatory skill assessments

1. RMSEs for water level, QC-ed currents only (speed and direction at all depth bins) (QC-ed currents available for Kill Van Kull, Jun-Jul 2021, Jan 2022).

2. Mean bias for water level, QC-ed currents only (speed at all depth bins) (QC-ed currents available for Kill Van Kull, Jun-Jul 2021, Jan 2022).

Optional skill assessments

3. RMSEs for temperature, salinity

4. Mean bias for temperature, salinity

5. Assessment of along channel currents and mean profile at Kill Van Kull (Calculate mean principal direction and project to that direction)

6. Assessment of current velocity time series using vector correlation (Kundu, 1976; Liu et al., 2020)

7. Ways to evaluate peak and trough timing, e.g. RMSE of difference in time of high/low water level between model and observation (THW-thw, TLW-tlw, where upper case is model and lower case is obs), correlation lags

8. RMSEs for water level, currents (speed and direction) RAW data & spin-up period.

9. Also consider other independent observations, e.g.
    a. XBT data from M/V Oleander: ERDDAP,
    b. satellite SST (and here)/SSS,
    c. HF Radar,
    d. underwater glider observations, e.g. IOOS Glider DAC: https://gliders.ioos.us/
    e. satellite-tracked surface drifter data

10. Process-based evaluations:
    a. Jan 1 - Mar 31, 2022: includes Nor'easter events (Jan 14-17; Jan 28-29)
    b. Jul 1 - Sep 30, 2021: includes hurricane events (Elsa, Henri, Ida)
    c. Example processes that can be studied: thermocline location, location of Hudson River Plume, and upwelling events, interaction with Hudson Canyon

4. SUCCESSES, CHALLENGES, AND LESSONS LEARNED

In the 2nd round of this UFS CAT model evaluation, the team became more cohesive and meetings within the two model groups (FVCOM and SCHISM) were very productive. The overall coordination and communication within the UFS CAT enabled the team to transition from the "storming and norming" phases into the "performing" phase, with the goal for advancing the models so that they are ready for operational applications within the UFS Coastal framework. It is important to note that the two models that were evaluated began at different readiness levels for various applications. More details on the starting point for readiness levels as well as a gap analysis will be discussed in the lessons learned below (Section 4.3). As such, the teams progressed at different paces in the development cycle. By understanding the desired operational objectives for the models, the UFS CAT was able to identify the current gaps for each model in order to advance the models through NOAA's readiness level evaluation for operational applications. As such, no specific results will be provided in this report and rather some results are included in publications that have resulted from this project. Instead, the discussion and summary will focus on the successes, challenges and lessons learned throughout this round.

4.1. Successes

<u>Success 1- training the next generation of ocean modelers</u>

As the UFS emphasizes that federal agencies need to collaborate on code development, it is critical to involve academic partners. As a result, the UFS CAT provides an opportunity to train the next generation of ocean modelers and build a sustainable workforce and research to operations (R2O) process. NOAA was able to provide ~$1.65M over the past three years to support academic partners. In total, nine postdocs and graduate students, as well as their faculty advisors, have been involved in the model evaluation and trained so far as part of this project. It was noted that the process of obtaining knowledge and skills through the model evaluation leads to collaborations among these tester groups and quicker progress for each tester. It was also noted that the effort enabled close partnership between the testers (students or postdocs) and their advisors, with the NOAA development and operational teams. This helps streamline the R2O process, facilitating the latest scientific advances into NOAA UFS operations.

<u>Success 2- learning NOAA's marine navigation model requirements and skill assessment</u>

Through the hindcast simulation and skill assessment process, testers were trained in coastal ocean modeling, as well as exposed to model simulations and working in an HPC environment. Thus, the model evaluation effort allows for communication of NOAA's marine navigation model requirements, and to match expectations on skill assessments. In addition, regional experts (e.g., Joseph Sienkiewicz, the NWS Ocean Prediction Center) contributed local knowledge (e.g., ice characteristics in Hudson River and flow downstream) to help interpret the model simulation results and skills, underscoring the importance of imparting local knowledge to development/improvement of operational models.

Success 3- conducting skill assessment

Testers were able to conduct skill assessment using multiple metrics and observational datasets. Testers, all new to FVCOM and SCHISM, were successful in conducting this skill assessment, and their simulations mostly met the NOS skill assessment acceptance criteria for water levels and currents.

They were exposed to observational datasets that they may not have been exposed to prior. This allowed them to develop their understanding of observational dataset limitations, and the importance of the quality control of those datasets before direct comparison with model output (e.g., take daily average of HF Radar currents instead of hourly snapshots to avoid noise). The testers used the traditional NOS skill metrics but a few used additional optional metrics that were found to be useful, especially for process studies.

Testers were encouraged to explore these optional skill assessment criteria using a variety of available *In Situ* observational datasets (see Section 3.5). Most testers were able to evaluate the modeled water temperature and salinity and compare it with buoy observations. One tester compared model results with XBT water temperature data and HF radar surface water currents for conducting model sensitivity tests (Park et al., in review). Underwater glider data were also used to assess water temperature and salinity vertical transects modeled by multiple simulations.

Required skill assessments included RMSE and mean bias calculations for water level and quality-controlled currents (speed and direction at all depth bins). Note that the quality-controlled currents for sufficiently long periods were available only for the Kill Van Kull area. Simulated water level results using mean RMSE across all stations, models, and testers mostly met the NOS marine navigation acceptable error of 0.15 m, and simulated current speed results using the same mean RMSE all met the NOS marine navigation acceptable error of 0.26 m/s. This success proved that both community-based models were accessible to testers with varying degrees of modeling experience and with a relatively short amount of time to go through the whole end-to-end process of testing and evaluating coastal ocean models. It also demonstrated these two models are potential candidates for the coastal component of the UFS with regards to skill.

Bathymetry data issues (see Section 3.3) were also identified as part of the skill assessment process, and these were reported and fixed during the iterative mesh development. This feedback loop was a success in helping to improve model mesh and performance (Zhang et al., 2024).

Success 4 - conferences and publications

Successes, challenges, and lessons from Round 1 were documented in an earlier report that was published as a NOAA Technical Memorandum (Seroka et al., 2024). In Round 2, a few other manuscripts were published and/or submitted for publication in peer-reviewed journals (e.g., Zhang et al., 2024; Park et al., in review), and 10 conference presentations were given, which are listed in Appendix A. The co-leads of the UFS CAT model evaluation also anticipate developing a short article for a peer-reviewed journal that highlights the overall collaborative efforts with this model evaluation.

The model testers and the co-leads actively presented the model evaluation efforts at conferences. Most notably, UFS CAT members that include Shachak Pe'eri, Greg Seroka, John Kelley, Ayumi Fujisaki-Manome, Saeed Moghimi, and Tracy Fanara co-chaired a session “Coastal Modeling and Evaluation for the Unified Forecast System (UFS) and Other Applications" at the 22nd Symposium on the Coastal Environment of the American Meteorological Society (AMS) annual meeting in Baltimore, MD on January 28-February 1, 2024. Multiple testers presented their research associated with this UFS CAT model evaluation effort (see below). A similar session proposal was accepted for the AMS 2025 annual meeting to continue to highlight the team's achievements. In addition, several topics associated with this effort were presented at the Unifying Innovations in Forecasting Capabilities Workshop 2024 (UIFCW24) in Jackson, Mississippi from July 22-26, 2024. There were additional presentations by the team members, including the overview of the UFS CAT model evaluation effort at the 2023 Hydrographic Survey Review Panel Fall meeting and a poster presentation at the Annual Partners Meeting of the Cooperative Institute for Great Lakes Research (May 2024), and at the NOS leadership team visit to NOAA Great Lakes Environmental Research Laboratory on September 23, 2024, prior to the Hydrographic Services Review Panel (NOS Federal Advisory Committee) in Detroit.

4.2. Challenges

Challenge 1- Consistency

As a goal of this model evaluation effort is to expose testers to NOAA operational models and train this next generation of coastal modelers, flexibility and trial and error were encouraged so that testers learn along the way. This led the project to allow for each tester to have their own baseline configuration rather than one baseline configuration for all testers of each model. As a result, each tester conducted their own “baseline” configured simulation, which acted as a starting point for running other model configurations, such as different atmospheric forcing data, air-sea flux formulation, hydrology inputs, etc. Testers were encouraged to follow guidelines as much as possible, to allow for sharing knowledge and lessons learned among the testers.

The multiple “baselines” introduced a challenge for testers to discuss issues they face and to troubleshoot problems since every tester used different model setups. These issues encountered included overheating in shallow areas, high velocities along the boundary, etc. A few testers overcame this problem by sharing their experience and applying the remedies, demonstrating initiative and developing leadership qualities. Another remedy for testers to address issues was through assistance from the model developers, who shared information about available learning resources with testers (an FVCOM online class by Dr. Chen, which had its first session in spring semester 2024; FVCOM 2024; and recordings of the 2024 SCHISM course by Dr. Zhang; SCHISM 2024). Testers were also encouraged to attend the NOAA annual workshops for SCHISM (July 19-20, 2023) and FVCOM (August 8-9, 2023) to learn from the experiences of other ocean modelers.

While the above remedies had helped testers in addressing their modeling issues, it was agreed upon that there still needed to be more specific guidance in addressing similar issues over the remaining months of the project. For example, it was suggested to provide a specific benchmark case as a reference working model configuration. This includes the specific version of the model core to run and how to set the namelist to make the model work for the specific configuration.

Challenge 2- Issues with the DEMs

Because DEM is the foundation of hydrodynamic model simulations, any issues with DEM will generate significant model simulation errors and uncertainties. Substantial time, attention, and efforts are needed to 1) create the DEM, 2) identify and validate potential issues with DEM through model simulations, 3) fixing the DEM issues identified, and 4) making sure that the corrected DEM fixed the problem in the model simulations.

It can be difficult and time-consuming to identify issues in the DEM without moving through the entire model set up, simulation, and skill assessment process, and then repeating that process once the updated bathymetry is received. This is a very time consuming but important effort for the testers as well as the DEM data providers. As stated in Section 3.3, bathymetry data issues were identified and reported by the testers during their evaluation process. For example, a tester evaluated modeled currents against observations at Kill Van Kull and found model accuracy issues at that location. This helped reveal issues in the DEM nearby. The two DEM issues that were identified in this model evaluation process were 1) the disconnected river in the DEM over the Arthur Kill area (Figure 3a), and 2) inconsistent DEM across Arthur Kill (Figure 4a) which led to a significant jump in the bathymetry depths in that channel.

These issues were remedied in the latest version of the NBS and the updated bathymetry data was provided to testers (Figure 3b, 4b). These DEM updates benefited not only the testers in their model skill but also other users of NBS data. The process of identifying, reporting, and fixing bathymetry data issues was again a critical part of iterative mesh development, underscoring the importance of identifying DEM issues early on in the mesh development as it impacts the model performance across all rounds; however, as stated earlier, it is not an easy task to identify these issues without going through the whole process.

A potential solution that could be implemented in the future to help streamline the DEM process and save time and effort is to utilize all the information in NBS in preparing the model's DEM, e.g. not just the bathymetry depths but also the information about source and uncertainty. While a more ad hoc approach to generate the DEMs was used in this project, NBS has implemented a validated process to generate the best available DEM which should reduce uncertainties in the whole DEM process. NBS presently covers the U.S. East Coast and Gulf of America but is looking to cover all U.S. coasts in future years, which then will provide modelers access to best available DEMs wherever they need to run a coastal simulation.

Challenge 3- Needs for local knowledge to interpret the results

Model errors were often difficult to assess without local oceanographic/hydrodynamic knowledge. For example, cold bias in water temperature was noted in the estuary, which appeared to be impacted by lack of representation of warm water point sources by power plants (Georgas, 2010) and/or sea/river ice. Another example is the judgment not to touch jumps in the bathymetry data when the jump could actually be real. Local knowledge and experience can significantly help in determining the real-world features. The importance of local knowledge also applies to understanding the model-observation differences, such as those from underwater gliders (e.g., subsurface warm core rings impinging on the shelf), HF radar (e.g., valid spatial coverage for the data), and XBTs.

4.3. Lessons Learned

Based on the success and challenges mentioned above, multiple lessons learned will be implemented in the next rounds. The categories of lessons learned are: gap analysis, matching expectations, and update to skill assessment software.

Lesson 1- Gap analysis

Through the model evaluation effort, issues inherent to each model were identified based on each tester's effort. This led the team to establish the goal of analyzing the gaps in each model's performance and skill and addressing those gaps so the two models end up on a level "playing field" and more ready for NOAA operations. The result is having multiple models ready for operations rather than one which can be a single point of failure risk. However, as Round 2 is in the early stages of the project (3D with atmospheric forcing, without any coupling), only minor gaps have been identified so far. We anticipate discovering and documenting gaps specific to each model within the coupled model ecosystem in Rounds 3 and 4.

Some of these minor operational gaps identified include overheating in shallow water due to issues in atmospheric forcing resolution, and maintaining stratification in the summer due to forcing issues and/or vertical discretization in the model. Addressing these issues often require model-specific guidance (often from the model developer or more experienced testers) on how to set up the model to resolve the issue. This collaboration among model developers and experienced and less experienced testers builds the community supporting each model, which is a critical component of the community model being ready for Earth Prediction Innovation Center (EPIC)-supported NOAA UFS operations.

Lesson 2- Realistic expectations

Another lesson learned was the co-leads being realistic in expectations that the testers did not have the capacity and time to consider both accuracy and computational efficiency in the early stages of the project. At the start of the model evaluation process, testers and project co-leads set out to achieve multiple requirements, including accuracy and computational efficiency. In the first couple of rounds, however, testers have focused on setting up and learning about model environments and configurations, with less of a focus on computational efficiency. This required co-leads to adjust accordingly and be realistic in their expectations about what can be accomplished.

Now that the testers have "gotten their feet wet", the balance between accurate, high resolution simulations and practical constraints with computational resources should be further pursued in the following rounds. While better horizontal and vertical resolutions will lead to more accurate representations of physical processes that are important for the marine navigation requirements, such high resolution configurations will lead to high computational expense, which is not practical for operational environments at NOAA.

Lesson 3 - Update to skill assessment software

The existing NOS Fortran-based skill assessment code package was made available to the testers but not utilized due to various reasons. This further motivated the ongoing effort of NOS's development of a next generation skill assessment software for coastal ocean models. This package will likely incorporate new skill metrics with the traditional ones, and also allow for comparison to different observational datasets, especially subsurface and spatial, to have more consistency and usability, and allow for process-based skill assessment. This software will eventually be modified to meet the UFS METplus requirement.

5. NEXT STEPS AND RECOMMENDATIONS

The model evaluation is already in the third and fourth rounds that will incorporate wave processes and testing UFS-Coastal (Round 3), as well as evaluating National Water Model (NWM) outputs as river forcing (Round 4). The co-leads will provide wave (WAVEWATCH III) forcing data as well as the UFS-Coastal code base (Round 3), and NWM forcing data (Round 4) for the testing. The results from the tests will be provided in future reports for Rounds Three and Four.

ACKNOWLEDGEMENTS

The UFS CAT team would like to thank the **NOAA NOS Water Team** for funding support for two years of research (FY23 and FY24), Cooperative Institute for Marine and Atmospheric Research (CIMAR), Cooperative Institute for Great Lakes Research (CIGLR), Mid-Atlantic Regional Association Coastal Ocean Observing System (MARACOOS), Southeast Coastal Ocean Observing Regional Association (SECOORA), and Great Lakes Observing System (GLOS). We would also like to thank **Dr. Changsheng Chen** and **Dr. Jianhua Qi** from School of Marine Science and Technology (University of Massachusetts – Dartmouth) and **Dr. Joseph Zhang** from Virginia Institute of Marine Sciences (William & Mary) for their support throughout the model evaluation process, and educating and engaging with testers from government and academia. Similarly, we would like to thank **UT Austin/TACC** for providing development work space and IT support for developers and testers to evaluate the models. Finally, we would like to thank **Drs. Ufuk Turunçoğlu**, **Panagiotis Velissariou**, **Yunfang Sun**, and **Jana Haddad** who are the core developer team working with representatives from the coastal ocean modeling community to establish the first iteration of the UFS-Coastal code base infrastructure being utilized in the next few rounds of the evaluation work.

REFERENCES

Amante, C.J.; Love, M.; Carignan, K.; Sutherland, M.G.; MacFerrin, M.; Lim, E. , (2023). Continuously Updated Digital Elevation Models (CUDEMs) to Support Coastal Inundation Modeling. Remote Sens., 15, 1702. https://doi.org/10.3390/rs15061702

Aquaveo. 2021. SMS 13.2. US Army Corps of Engineers. https://www.aquaveo.com/software/sms-surface-water-modeling-system-introduction

Militello, A. and Alan K. Zundel, A., 1999. Surface-Water Modeling System Tidal Constituents Toolbox for ADCIRC, Coastal Engineering Technical Note IV-21. https://erdc-library.erdc.dren.mil/jspui/bitstream/11681/2241/1/CETN-IV-21.pdf

Beardsley, R.C., & Chen, C. (2014). Northeast Coastal Ocean Forecast System (NECOFS): A Multi-scale Global-Regional-Estuarine FVCOM Model. American Geophysical Union, Fall Meeting 2014, abstract id. OS23C-1211

Bureau of Transportation Statistics (BTS), https://www.bts.gov/content/tonnage-top-50-us-water-ports-ranked-total-tons, last accessed on March 18, 2025

Chen, C. H. Liu, R. C. Beardsley, 2002. An unstructured, finite-volume, three-dimensional, primitive equation ocean model: application to coastal ocean and estuaries. Journal of Atmospheric and Oceanic Technology, 20, 159-186.

Chen, C., Beardsley, R. C., & Cowles, G. (2006). An unstructured grid, finite volume coastal ocean model (FVCOM) system. Oceanography, 19(1), 78–89. https://doi.org/10.5670/oceanog.2006.92

Egbert, Gary D., and Erofeeva, Svetlana Y., (2002). "Efficient inverse modeling of barotropic ocean tides." Journal of Atmospheric and Oceanic Technology 19.2:183-204. https://www.tpxo.net/tpxo-products-and-registration

Foreman, M. G. G., Henry, R. F., Walters, R. A., and Ballantyne, V. A. (1993), A finite element model for tides and resonance along the north coast of British Columbia, J. Geophys. Res., 98(C2), 2509–2531, doi:10.1029/92JC02470.

FVCOM 2024 Online Training Course, https://www.fvcom.org/?p=13, last accessed on October 16, 2024.

Gardner, J.V., Mayer, L. A., and Armstrong, A. A., (2005). “U.S. Law of the Sea Mapping”, Hydro International, vol. 9. Geomatics Information & Trading Center - GITC, Lemmers, Amsterdam, The Netherlands, pp. 42-45.

Georgas, N., (2010), “Establishing Confidence in Marine Forecast Systems: The Design of a High Fidelity Marine Forecast Model for the NY/NJ Harbor Estuary and its Adjoining Coastal Waters”, 10.13140/RG.2.2.14602.98240.

Huang, W., Zhang, Y.J., Wang, Z. et al., (2022).Tidal simulation revisited. Ocean Dynamics 72, 187–205. https://doi.org/10.1007/s10236-022-01498-9

Huang, M., Arcas, D., Van der Westhuysen, A., Seaton, C., Warner, J., Moghimi, S., Peeri, S., Wei, Y., Yang, Z., Haas, K., Taylor, A., Pringle, W., Zyserman, J., (2022). Risk Reduction Sub-Group Models Model Evaluation Report. https://doi.org/10.25923/x2jk-xh48

Kundu, J.K., (1976). Ekman veering observed near the ocean bottom. Journal of Physical Oceanography, 6, 238-242.

Liu, Y., Weisberg, R.H., Zheng, L., (2020).: Impacts of hurricane Irma on the circulation and transport in Florida Bay and the Charlotte Harbor estuary. Estuaries and Coasts, 43, 1194-1216, https://doi.org/10.1007/s12237-019-00647-6

Lyard, F. H., Allain, D. J., Cancet, M., Carrère, L., and Picot, N., (2021). FES2014 global ocean tide atlas: design and performance, Ocean Sci., 17, 615–649, https://doi.org/10.5194/os-17-615-2021 .

NOAA, 2022 - UFS Coastal Applications Team - Water Quantity Marine Navigation, Sub-Application Tiger Team: Report, NOAA Technical Memorandum (NOS, 34; NWS 02; OAR, 02), Silver Spring, MD, June, 2022, 27 pp. https://repository.library.noaa.gov/view/noaa/60886

NOAA, 2025a, NAO 216-115B: Research and Development in NOAA, https://www.noaa.gov/organization/administration/nao-216-115b-research-and-development-in-noaa, last accessed on March 15, 2025.

NOAA, 2025b, NAO 216-105B: POLICY ON RESEARCH AND DEVELOPMENT TRANSITIONS, https://www.noaa.gov/organization/administration/nao-216-105b-policy-on-research-and-development-transitions, last accessed on March 15, 2025.

Mani, Soroosh et al., (2021). OCSMesh: a data-driven automated unstructured mesh generation software for coastal ocean modeling. https://doi.org/10.25923/csba-m072

Park, K., Zhang, J., Di Lorenzo, E., Seroka, G., Fujisaki-Manome, A., Pe'eri, S., Moghimi, S., and Kelley, J., Evaluation of a 3D Unstructured Grid Model for the New York-New Jersey Harbor under Different Forcing Sources, *Ocean Modeling*, in review.

Roberts, K. J., Pringle, W. J., and Westerink, J. J., (2019). OceanMesh2D 1.0: MATLAB-based software for two-dimensional unstructured mesh generation in coastal ocean modeling, Geosci. Model Dev., 12, 1847–1868, https://doi.org/10.5194/gmd-12-1847-2019

SCHISM 2024 course at the Virginia Institute of Marine Science (2024), recordings of the 13-week course is available at https://ccrm.vims.edu/w/index.php/Online_tutorials#13-week_SCHISM_course_at_VIMS_.282024.29.

Seroka, G., Kelley, J., Sienkiewicz, J., Seaton, C., Anderson, E.J., Allen, A., Forbes, C., Fujisaki-Manome, A., Haas, K.A., Lindley, C., Roman, D., and Weston, A. (2022), Unified Forecast System (UFS) Coastal Applications Team - Water Quantity Marine Navigation Sub-Application Tiger Team: Report, NOAA Technical Memorandum, NOS(34), 1–27. https://doi.org/https://doi.org/10.25923/x7zr-nj33.

Seroka, G., Fujisaki-Manome, A,, Kelley, J., Pe'eri, S., Sienkiewicz, J., Feyen, J., Doty, O., Ide, K., Gramp, B., Ogden, F., Fanara, T., Myers, E., Moghimi, S., Cockerill, T., Wu, W., Anderson, E., Huelse, K., Forbes, C., Liu, Y., John, S., Di Lorenzo, E., Park, K., Wipperfurth, S., Sannikova, N., Titov, V., Wei, Y., Akan, C., Mani, S., Lindley, C. (2024), UFS Coastal Applications Team Report Round 1 Summary of a Unified Forecast System Model Evaluation for Marine Navigation, NOAA Technical Memorandum NOS 36/NWS04/OAR04, May 2024, https://doi.org/10.25923/nws6-kx30.

Tuckman, Bruce W (1965). "Developmental sequence in small groups". *Psychological Bulletin*. **63** (6): 384–399. doi:10.1037/h0022100

UFS (2019). Unified Forecast System Overview. Available at https://ufscommunity.org/wp-content/uploads/2019/10/201903xx_UFS_Overview.pdf

Yang, Zizang et al. (2019). NOAA's Gulf of Maine Operational Forecast System (GOMOFS): Model Development and Hindcast Skill Assessment. https://doi.org/10.25923/0m2e-xg81

Van der Westhuysen, A., Ogden, F., Flowers, T., Fanara T., Myers, E., Dean, C., Allen, A., Lindley, C., Zachery, B., Fujisaki-Manome, A., Warner, J., Doran, K., Palmsten, M., Massey, C., Sanders, B., and He., R., Whitepaper on the development of a Unified Forecast System for Coastal Total Water Level prediction, Unified Forecast System (UFS) Coastal Application Team, Water Quantity Total Water Level Prediction Sub-application, NOAA Technical Memorandum NOS, 33, 1–60. https://doi.org/https://doi.org/10.25923/170a-cv35.

Wyllie, K. and Rice G., (2020). Building the National Bathymetry, published on February 07, 2020. https://nauticalcharts.noaa.gov/updates/building-the-national-bathymetry/

Zhang, Y., Ye, F., Stanev, E.V., Grashorn, S. (2016) Seamless cross-scale modeling with SCHISM, Ocean Modelling, 102, 64-81.

Zhang, J., (2023a). SCHISM Online Tutorial on January 25, 2023: Mesh generation training for beginners, http://ccrm.vims.edu/yinglong/wiki_files/Presentation_GridGen_SCHISM_Joseph_Zhang_01252023.mp4, last accessed on August 31, 2023

Zhang, J., (2023b). SCHISM Online Tutorial on July 19, 2023: Mesh generation training for beginners, Two part series available at http://ccrm.vims.edu/w/index.php/Online_tutorials, last accessed on August 31, 2023

Zhang, J., Anderson, J., Park, K., Wu, C.H., Wipperfurth, S., Anderson, E., Pe'eri, S., Beletsky, D., Titze, D., Di Lorenzo, E., Moghimi, S., Seroka, G., Myers, E., Fujisaki-Manome, A., Kelley, J. (2024), Debunking common myths in coastal circulation modeling, *Ocean Modelling*, 190, https://doi.org/10.1016/j.ocemod.2024.102401.

APPENDIX A - UFS CAT PRESENTATIONS RELATED TO ROUND 2

Presentations at the 22nd Symposium on the Coastal Environment of the American Meteorological Society (AMS) Annual Meeting in Baltimore, MD for January 28-February 1, 2024.

- Wu, W., Myers, E., Chen, C., Li, S., Qi, J., Xu, Q., Xu, J., Zhang, A., Seroka, G., Fujisaki-Manome, A.“Leverage NOAA NECOFS-Based OFS development for Supporting the UFS CAT Project”
- Seroka, G., Wang, Y-M., Hardy, R., Ahlgren, K., Pe’eri, S., Tang, L., Myers, E., Zhang, J., “Marine Geoid Validation Using Ocean Modeling Sea Surface Topography”
- Sannikova, N., Y. Wei and V. Titov, “Evaluation of SCHISM Model for the Unified Forecast System: Procedures, Results and Challenges”
- Doty, O., Fujisaki-Manome, A., and Park, K., “River Sensitivity in Coastal Modeling with SCHISM”
- Gramp, B.E., K. Ide, A. Fujisaki, and G. Seroka, “Assessing Water Level Hindcasts from FVCOM in the New York Harbor Area”
- Pereira, K., E. J. Anderson, A. Fujisaki-Manome, and J. Kessler. “Incorporating Satellite Ice Data for Enhanced Operational Forecasting : Bridging Ice Floe Trends and Dynamic Parameterization”
- Fujisaki-Manome, A., Hu, H. Wang, J. Westerink, J., Wirasaet, D., Ling G, Choi, M., Moghimi, S., Myers, E., Abdolali , A., Dawson C., Janzen. C., "Advanced sea ice modeling for integration into a storm-surge, wave, and ice forecasting system for Alaska's coasts"

Other presentations

- Haddad, J., Seroka, G., Moghimi, S., Kelley, J., Fujisaki-Manome, A., Doty, O., Turuncoglu, U., Tsay, A., Huang, M., Velissariou, P., Sun, Y., Khazaei, B., Abed-Elmdoust, A., Kurapov, A., Myers, E., Allen, C., Pe’eri, S., Fanara, T., Snowden, D., Burke, P., “User support, external testing, and project planning of the UFS Coastal coupling infrastructure in partnership with UFS Coastal Applications Team”, the Unifying Innovations in Forecasting Capabilities Workshop 2024 (UIFCW24) in Jackson, Mississippi from July 22-26, 2024.
- Fujisaki-Manome, A., Seroka, G., Kelley, J., Pe’eri, S., “Unified Forecast System (UFS) Coastal Applications Team HSRP Update”, the Hydrographic Survey Review Panel Public Meeting, September 27-29, 2023.
- Doty, O., Fujisaki-Manome, A., Seroka., G., Pe’eri, S., and Kelley, J., “NOAA Unified Forecast System Coastal Applications Team (UFS CAT) Model Evaluation Overview”, Cooperative Institute for Great Lakes Research Annual Partners Meeting, May 6-7, 2024.