NOAA Technical Memorandum NOS 36
NOAA Technical Memorandum NWS 04
NOAA Technical Memorandum OAR 04

UFS Coastal Applications Team Report
Round 1 Summary of a Unified Forecast System Model Evaluation for Marine Navigation

Silver Spring, Maryland
May 2024

noaa National Oceanic and Atmospheric Administration

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

National Weather Service
Oceanic and Atmospheric Research

National Oceanic and Atmospheric Administration
U.S. Department of Commerce

The National Ocean Service (NOS) is one of six major divisions within the National Oceanic and Atmospheric Administration (NOAA), which is housed within the U.S. Department of Commerce. NOS is the nation's most comprehensive coastal agency with world-class expertise in science, technology, and management. Although NOS delivers a diverse suite of products and programs, the main mission areas are reflected in the three primary sections of its budget: Navigation, Observations, and Positioning; Coastal Science and Assessment; Ocean and Coastal Management and Services.

The National Weather Service (NWS) has played a key role in protecting American lives and properties for over a century. The timely provision of reliable weather, water, climate, and environmental information has supported the Nation's social and economic development. NWS offices in communities across the United States and its territories, supported by regional and national centers, provide the authoritative information needed by Americans, including national, regional, state, tribal, and local authorities, to plan, prepare, mitigate, and respond to natural and human-caused events.

Oceanic and Atmospheric Research (OAR) is a division of the National Oceanic and Atmospheric Administration (NOAA). OAR is also referred to as NOAA Research. NOAA Research is the research and development arm of NOAA and is the driving force behind NOAA environmental products and services aimed at protecting life and property and promoting sustainable economic growth. Research conducted by programs within NOAA and through collaborations outside NOAA, focuses on enhancing the understanding of environmental phenomena such as tornados, hurricanes, climate variability, changes in the ozone layer, El Niño/La Niña events, fisheries productivity, ocean currents, deep sea thermal vents, and coastal ecosystem health.

NOAA Technical Memorandum NOS 36
NOAA Technical Memorandum NWS 04
NOAA Technical Memorandum OAR 04

UFS Coastal Applications Team Report
Round 1 Summary of a Unified Forecast System Model Evaluation for Marine Navigation

Greg Seroka[1] (Team Co-Lead), Ayumi Fujisaki-Manome[2,3] (Team Co-Lead), John Kelley[1] (Team Co-Lead), Shachak Pe'eri[4] (NOS Co-Lead), Joe Sienkiewicz[5] (NWS Co-Lead), Jesse Feyen[3] (OAR Co-Lead), Olivia Doty[2], Kayo Ide[6], Brendan Gramp[6], Fred Ogden[7], Tracy Fanara[8], Edward Myers[1], Saeed Moghimi[1], Timothy Cockerill[9], Wei Wu[1], Eric Anderson[10], Kaitlin Huelse[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[1,20], Carolyn Lindley[21]

1. NOAA/NOS/OCS
2. University of Michigan/CIGLR
3. NOAA/OAR/GLERL
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, University of Hawaii
17. NOAA/OAR/PMEL
18. University of Washington
19. University of North Florida
20. SFI
21. NOAA/NOS/CO-OPS

May 2024

noaa **National Oceanic and Atmospheric Administration**

U. S. DEPARTMENT OF COMMERCE
Gina Raimondo, Secretary

National Oceanic and Atmospheric Administration
Richard Spinrad, Under Secretary

National Ocean Service
Nicole LeBoeuf, Assistant Administrator

Office of Coast Survey
Rear Admiral Benjamin Evans
Director

Coast Survey Development Laboratory
Corey Allen
Acting Division Chief

NOTICE

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

The Unified Forecast System (UFS) Coastal Application Team (CAT) has developed model evaluation recommendations for selecting NOAA's next-generation numerical oceanographic circulation prediction models that will support safe and efficient navigation and provide improved marine forecasts. Based on the requirements identified in the Water Quantity Marine Navigation, Sub-Application Tiger Team Report (NOAA, 2022) and the operational requirements to operate within NOAA using NOAA IT infrastructure, two models were selected for evaluation: 1) Finite Volume Community Ocean Model (FVCOM); and 2) Semi-implicit Cross-scale Hydroscience Integrated System Model (SCHISM). Additional outcomes expected from this government-academic collaboration include: 1) building the scientific oceanographic community (next generation of coastal modelers), 2) exposing the ocean modeling community to NOAA operational requirements, and 3) testing models and the National Unified Operational Prediction Capability (NUOPC) coupling infrastructure.

The model developers of FVCOM (Dr. Chen from University of Massachusetts, Dartmouth) and SCHISM (Dr. Zhang from William & Mary, Virginia Institute of Marine Science) worked with the UFS CAT, and National Science Foundation (NSF) Texas Advanced Computing Center (TACC) to provide a test environment to configure the circulation models. Testers involved in model evaluation included faculty members and students from seven universities, Full-Time Employees (FTEs) and contractors from NOAA.

Candidate models were configured and the compilation was optimized to a reliable performance state. Then the testers were onboarded and standard test runs were performed for a given region and a given specified computational time / resources to allow for fair qualitative/quantitative model evaluation.

This report described the successes, challenges and lessons learned of the first round of ~~the~~ model evaluation, where the goal was team building and matching expectations of the testers. This is mainly because the testers came with different academic backgrounds in oceanography and varying experience with the two models that were selected for the model evaluation. The key steps of the model evaluation include:

- Generate mesh (or adjust mesh) based on available gridded bathymetry
- Conduct baseline simulations
- Compare model results with observations
- Explore additional results to each tester's baseline simulation
- Conduct skill assessment based on NOAA's model evaluation guidance

The outcomes from the first round of model evaluation (tide only) will feed into a second round that will focus on incorporating the atmospheric forcing components onto the existing tidal model, and switch to running the model in layered density (3D) configuration.

1. BACKGROUND

The Unified Forecast System (UFS) Coastal Application Team (CAT) is part of a larger development within NOAA, the U.S. Coast Guard (USCG), and academic partners to review NOAA needs and consolidate them into individual modeling systems (*i.e*., global and regional atmosphere, ocean, land, *etc*.) using a smaller set of coupled Earth System models that would continue to serve its various stakeholders (UFS, 2019). Under the UFS CAT Marine Navigation working group (also known under the title of "Safe, Efficient Navigation"), the goal is to evaluate the leading oceanographic circulation models that operate in complex coastal environments. The evaluation encompasses the whole research to operations (R2O) cycle for circulation models at NOAA's National Ocean Service (NOS), and is not limited to the accuracy of a model under certain conditions. Two models were selected for evaluation of the Marine Navigation working group: 1) Finite Volume Community Ocean Model (FVCOM) that is currently used in the National Ocean Service (NOS) Great Lakes forecast systems and Northern Gulf of Mexico forecast system and is under development for some Atlantic and Pacific Regional models; and 2) Semi-implicit Cross-scale Hydroscience Integrated System Model (SCHISM) that is currently used for 3D surge and tide models in the Atlantic Ocean (including Gulf of Mexico and Caribbean Sea) and is also being applied in the development of the Pacific 3D surge and tide model.

NOAA's National Ocean Service (NOS) in partnership with NOAA's Oceanic and Atmospheric Research (OAR), National Weather Service (NWS) and the coastal ocean modeling community develops its next generation coastal ocean coupling infrastructure for integration into NOAA's Unified Forecast System portfolio. This is in parallel with the UFS CAT evaluation efforts described in this report. As a result of the above partnership, the UFS-Coastal (ufs-coastal-model) coupling infrastructure and its downstream applications (ufs-coastal-apps) will be implemented and utilized by the evaluation team. Figure 1 shows the envisioned capabilities and general structure of the ufs-coastal.

The development and implementation steps 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.
- Utilizing the ocean modeling GitHub organization as the community platform to jointly develop the ufs-coastal-model and related applications (*e.g.*, marine navigation, risk reduction and total water level; see: https://github.com/oceanmodeling/ufs-coastal).
- Establishing a Continuous Integration and Continuous Delivery (CI/CD) pipeline to employ software development best practices.

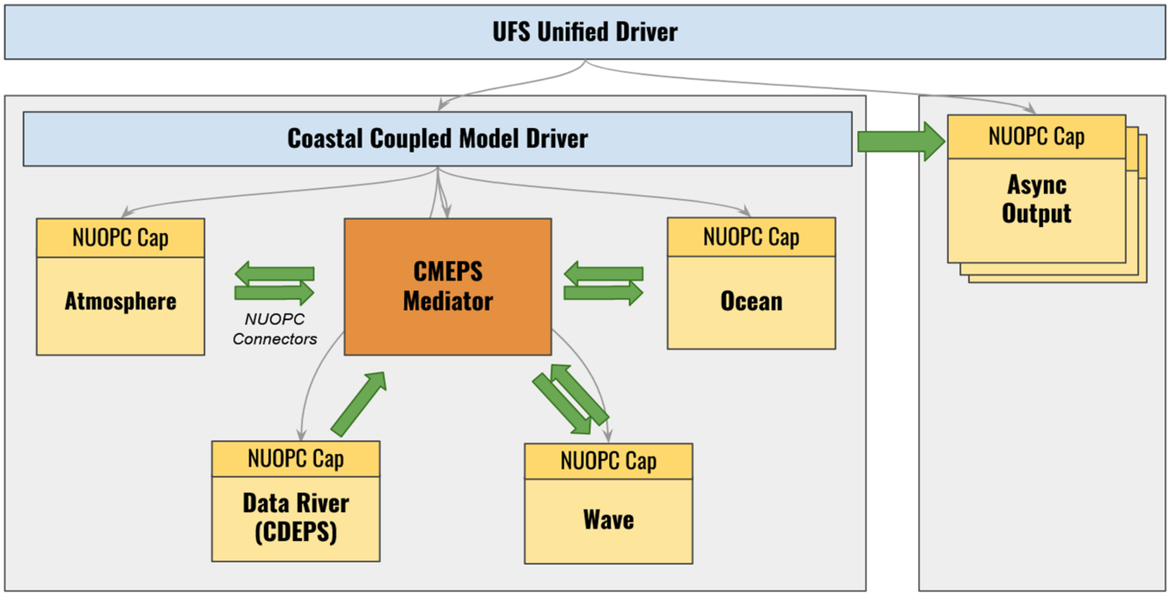

Figure 1. Envisioned UFS-Coastal coupling code infrastructure diagram

The model evaluation is conducted in three rounds.

- **Round 1 -** The models are evaluated in average water density (2D) mode and are not coupled to other models, where the water level (tidal) components were evaluated. The goal of Round 1 is to properly install the circulation models on the NSF TACC environment (with the help of the model developers), to make and evaluate tidal simulations to ensure quality 3D model simulations in subsequent rounds, and to get the testers familiar with the models. It is important to note that the testers in the model evaluation come from different universities with varying experience with the two selected models. This round allowed the testers to learn how to develop model meshes, how to prepare, make, and evaluate tidal simulations, and to transition from “forming” to “storming” and “norming” phases during this team development cycle (Tuckman, 1965).

- **Round 2 -** The models are evaluated in a layered water density (3D) mode, and are coupled with atmospheric models (GFS/HRRR). In addition to water levels, surface currents (top 4.5 m water layer), water temperature and salinity are also evaluated in the second round. The evaluation results will be shared with the developers and with the UFS community for feedback on performance time (run time), ease of operation, updating of elevation models, and calibration and skill assessment comparing model results with observations.

- **Round 3 -** The models are evaluated in (3D) mode and are coupled with atmospheric models (GFS/HRRR) and wave models (WaveWatch III). Similar to Round 2, water levels, surface currents (top 4.5 m water layer), water temperature and salinity are evaluated, and the evaluation results will be shared with the developers and with the UFS community for feedback on performance time (run time), ease of operation, updating of elevation models, and calibration and skill assessment comparing model results with observations.

Another goal of this UFS CAT effort is to share and include NOAA's model evaluation procedure and improve collaboration relationships between NOAA and academic partners by assessing model candidates together with the community. Furthermore, NOAA has recently launched the Earth Prediction Innovation Center (EPIC), which will be able to facilitate community collaboration by making NOAA's UFS models and its operational infrastructure available on a cloud platform for experimentation and improvement. In round 1, the key steps of the model evaluation include:

- Generate Mesh (or adjust the mesh) based on available gridded bathymetry
- Conduct baseline simulations
- Compare model results with observations
- Explore additional results to each tester's baseline simulation
- Conduct skill assessment based on NOAA's model evaluation guidance

2. MODEL EVALUATION REQUIREMENTS

Following UFS CAT Marine Navigation Water Quantity white paper (NOAA, 2022), the model evaluation requirements (Table 1) are based on communities, 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, and U.S. Coast Guard, 2022)

Forecast configuration	
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 up to 6 minutes
Forecast range	5 to 7 days, 14 days for planning (monthly/seasonal for lake/sea ice)
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 International Terrestrial Reference Frame (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, and U.S. Coast Guard, 2022)

	Accuracy (acceptable Root Mean Square Error)
Key user variables:	**Specifications**
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 - 0.25km/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. METHODOLOGY

3.1. Selection criteria

Two circulation models were selected based on requirements mentioned in the previous section and the following operational requirements:

- Finite Volume Community Ocean Model - FVCOM (Chen et al., 2002)
- Semi-implicit Cross-scale Hydroscience Integrated System Model - SCHISM (Zhang et al., 2016)

Both coastal ocean models utilize free surface 3-D primitive equations on unstructured grid frameworks. Both models are able to expand to a fully coupled current-ice-wave-sediment-ecosystem model with parallelization implemented. This type of 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.

The operational requirements that are also mentioned in UFS CAT Marine Navigation working group white paper (NOAA, 2022), 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 High Performance Computing (HPC) infrastructures are available at NOAA's NCEP Central Operations (NCO). Overall, the model robustness is a key requirement for operational use.

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. Operational grade models are expected to run reliably, and provide guidance for each cycle without fail. 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 requirement topobathy smoothing.

Code management - Even if NOAA is running the models operationally, it is very important that the code is managed 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 being coupled using the NOAA Environmental Modeling System (NEMS) 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 C00 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.

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

3.2. Team Development (setting up the models)

The UFS CAT appreciates the value in the developmental stage processes (Tuckman, 1965), and is interested in assisting the team to enter a stage consistent with the collaborative work put forth. It is important to transition from being in a comfort zone of non-threatening topics and risk the possibility of conflict (“Forming”), to "testing and proving" mentality (“Storming”) to a problem-solving mentality (“Norming”). Eventually, evolving into the team’s capacity, range, and depth of personal relations, expand 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.

Similar to any team development cycle, Round 1 in the UFS CAT model evaluation was to transition the testers from the “forming” phase into “storming” and “norming” phases. At first, the co-leads worked with the model developers, Dr. Chen from University of Massachusetts Dartmouth and Dr. Zhang from William & Mary (Virginia Institute of Marine Sciences), and NSF National Science Foundation (NSF) Texas Advanced Computing Center (TACC) to create a High Performance Computing (HPC) environment for each of the models. The goal of Round 1 was to properly install the circulation models on the NSF TACC environment (with the help of the model developers) and get the testers familiar with 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, Full-Time Employees (FTEs) and contractors from several groups at NOAA participated in the model evaluation. At first, access was provided to the model developers and power users 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 NSF/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 core hours for this start up project is 5000 core hours, with 40 compute nodes, 0.25 days of wall-clock time, up to 250 runs with 2500 core hours for each model. This HPC allocation is for an annual quota for all three rounds. The models are evaluated in average water density (2D) mode and are not coupled to other models, where the water level (tidal) components were evaluated.

In round 1, the key steps of the model evaluation included:

- Generate mesh (or adjust the mesh) based on available gridded bathymetry
- Conduct baseline simulations
- Compare model results with observations
- 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 was the New York Harbor, which 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. There are two main channels leading into the harbor, and several inland rivers connected to the bay. 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. Bathymetry data of the study site up to the continental shelf (Figure 2) was provided for mesh generation to all teams. NOAA's data repositories (OCS National Bathymetry Source (Wiley and Rice, 2020), and NCEI's Continuously Updated Digital Elevation Models (Amate et al., 2023) were used to download the data. All the dataset were vertically referenced to NAVD88 and horizontally to WGS-84. However, due to the data sources and the volume of data, different grid resolutions were provided based on the geographic coverage. The bathymetry grid resolution around the study area (New York Harbor, Long Island and its surrounding rivers) was provided at 4 meters, coastal areas around Rhode Island, Massachusetts and New Jersey were provided at 8 m resolution, offshore area on the continental shelf were provided at 16 m and 64 m, and the continental shelf and slope were provided at 100 m using the US Extended Continental Shelf program (Gardner et al., 2005).

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 due to lack of access to mesh generation tools (*e.g.*, SMS) and thus the ability to readily refine the mesh. Also, in some locations, the quality of the digital elevation models (DEMs) did not allow generating a mesh at a higher resolution. The teams instead 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. Also, the meshes can be updated and refined in Round 2 or Round 3.

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. All teams were provided access to the Texas Advanced Computing Center (TACC) Frontera system for testing. The number of nodes for each team's mesh ranged from 21,623 to 160,000.

The tools for creating mesh from bathymetry were the Surface-water Modeling System or SMS (Militello and Zundel, 1999; Aquaveo, 2021), OCSMesh (Mani, 2021), and OceanMesh2D (Roberts, 2019).

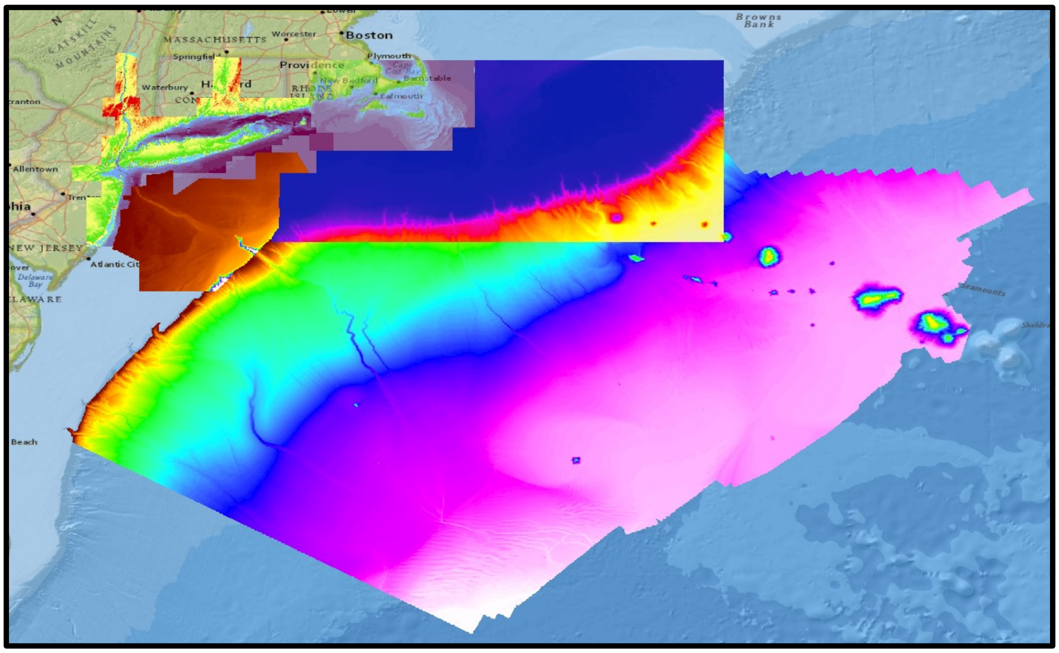

Figure 2. Available bathymetry datasets at the study site (NY Harbor) for model evaluation

3.4. Conduct baseline simulations

The source codes were both 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).

In round 1, the baseline simulation used open boundary conditions for evaluating only the tides. Most FVCOM testers used TPXO (Egbert and Erofeeva, 2002) as the boundary condition, where one team compared between TPXO, FES2014 (Lyard, 2021), and NECOFS (Beardsley, 2014).

The SCHISM tester teams used FES2014 as the boundary conditions in their model runs. Eight constituents were used for tidal forcing. Out of those eight constituents, the top four tidal constituents in the list below were evaluated in skill assessment (Yang et al. 2019):

- 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

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

1. Hindcast Period: 3 months, not including warm-up period.
    a. Jan 1 - Mar 31, 2022: includes Nor'easter (Jan 14-17; Jan 28-29)
    b. Jul 1 - Sep 30, 2021: includes hurricanes (Elsa, Henri, Ida)
2. Run time: Each 3-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.

Ideally, the testers would generate their own mesh, or at least go through that exercise. With the understanding that some may struggle creating an "ideal" mesh, a "Blessed" mesh by the developers was also provided to the testers as a backup (Figures 3 and 4). This back-up mesh was mainly for the testers who were limited in time and resource, and whose developed mesh did not work well in the subsequent evaluation processes.

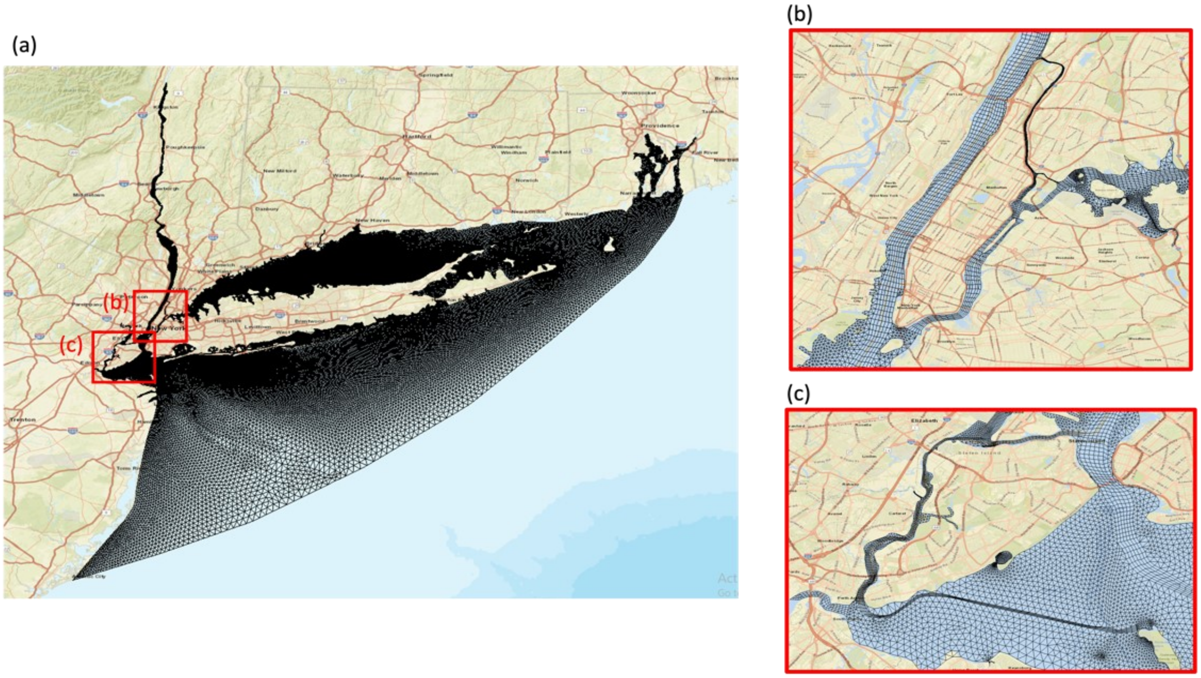

Figure 3. a) SCHISM mesh coverage ("blessed mesh") and subsets containing key navigation channels: b) Hudson and East Rivers around Manhattan (c) approach to NY Harbor passing from Sandy Hook to south Staten Island.

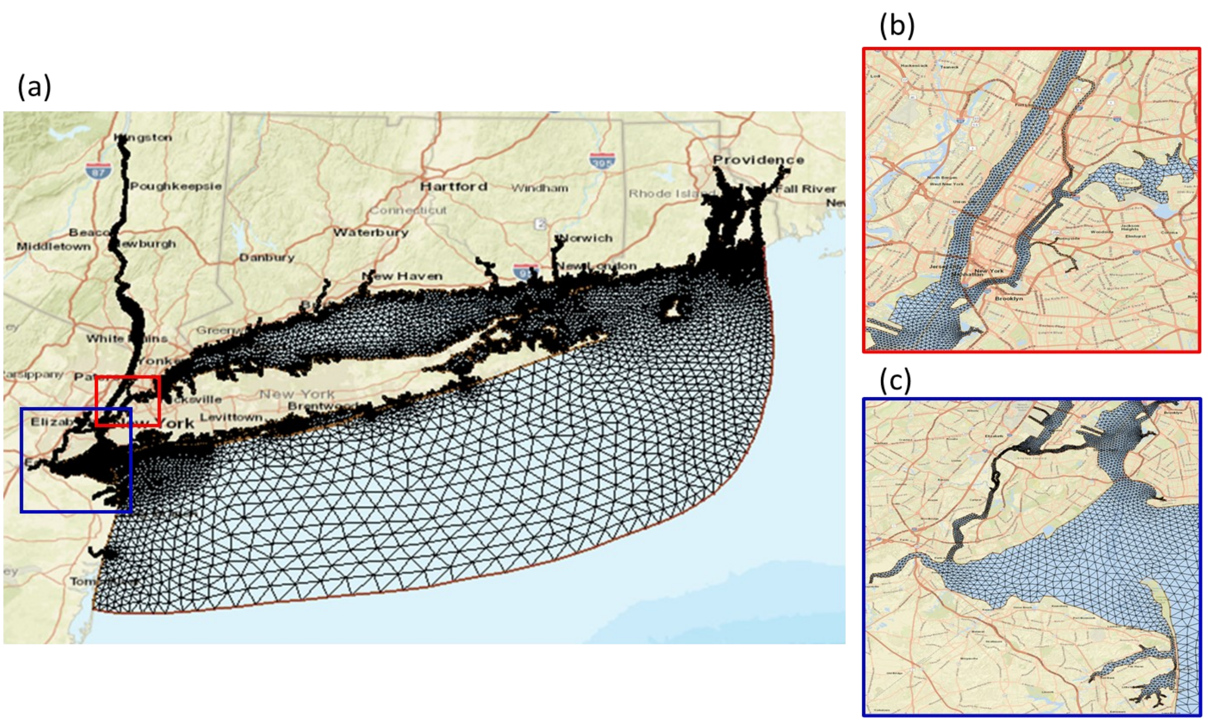

Figure 4. a) FVCOM mesh coverage ("blessed mesh") and subsets containing key navigation

channels: b) Hudson and East Rivers around Manhattan (c) approach to NY Harbor passing from Sandy Hook to south Staten Island.

3.5. Skill assessment using water level observations

The comparison of the model results utilized water level observations within the study area, and the observations were available to the testers on TACC. The testers used the evaluation metrics documented in the white paper from the Navigation working group (NOAA, 2022). The specific skill assessment criteria for Round 1 are as follows:

Evaluate the four major constituents (M2, S2, N2, and K1):

Method 1 (required) RMSE (in cm) for amplitude for each of the four constituents

RMSE (in degrees) for phase for each of the four constituents,

Method 2 (optional) Complex RMSE for each of the four constituents, computed from amplitudes and phases (avoids the wrap-around 360 degree issue for phases)

Some SCHISM testers used the equation from ‘Tidal simulation revisited - Huang et al. 2022’ (also see: Foreman et al. 1993) as follows:

$$Complex_{error} = \sqrt{((A_o \cos P_o - A_m \cos P_m)^2 + (A_o \sin P_o - A_m \sin P_m)^2)/2} \quad (1)$$

You can also see the script that uses the equation to calculate the complex RMSE in /scratch3/projects/CATUFS/KyungminPark/post-proc. The script is comp_TC.py in the path on Frontera.

4. SUCCESSES, CHALLENGES, AND LESSONS LEARNED

The main goals of Round 1 were getting the team familiarized with models and the development environment, and getting the testers to the “storming and norming” phases in the team development cycle. As such, no specific results will be provided in this report. Instead, the discussion on the results will focus on the successes, challenges and lessons learned throughout this round.

4.1. Successes

Success 1- communication and team building

Monthly meetings were conducted in order to share and communicate evaluation efforts between the testers. Separate meetings were held monthly for each model that was evaluated (*i.e.*, FVCOM and SCHISM), and included presentations to drive discussion and feedback from the model developers and other testers. Given that many of the students and government contractors were newer testers and are unfamiliar with the two models evaluated in this effort, they were also encouraged to attend a bi-weekly meeting hosted by a fellow student/contractor to prompt a lower-stakes environment for collaboration and questions. With both of these opportunities for feedback and collaboration, new testers were able to get an understanding of how circulation models are configured for NOAA operations. The UFS CAT team maintained live documentation of suggestions and lessons from testers and model developers to guide future efforts and improve on each round of this effort.

Success 2- learning the models/ exposure to NOAA ops

The FVCOM and SCHISM models were tested over two different hindcast periods that were established by the UFS CAT: January, 2022 - March, 2022 and July, 2021 - September, 2021.The models were evaluated by comparing to observation data for those time periods with the root mean square error (RMSE) and the complex RMSE (CRMSE) of phase and amplitude. Four major tidal constituents (M2, S2, N2 and K1) were evaluated successfully. Testers were able to use a prepared script to process the statistics of these constituents that was provided to the testers, or the testers could generate their own scripts.

Success 3- conducting skill assessment

The model results from each team were compared against 23 stations in the New York Harbor and surrounding area. The observations were analyzed without any data filtering. The skill assessment produced a RMSE of less than 10 cm in amplitude for all cases and constituents. In both years of the simulation, the tidal constituent M2 (the principal lunar semidiurnal constituent) produced the highest amplitude errors throughout all testing groups, with the maximum being 14.7 cm in 2021. The phase results for each year were less consistent than the amplitude results, with RMSE values for K1, the lunar diurnal constituent, varying as high as 30.1 degrees in 2021. For the same time period, some groups achieved a phase RMSE as low as 4.3 degrees for the constituent M2, which had the smallest range of phase RMSE values between all teams.

4.2. Challenges

Challenge 1- consistency and guidance

While the overall model evaluation process has allowed testers freedom and flexibility in their configurations and methodology, it has been suggested that there should be more guidance in mesh coverage, resolution, boundary conditions, and so on. This feedback prompted the co-leads to develop a baseline configuration for testers to begin with, if they chose. This baseline also allows for consistent result comparison amongst the testing groups, if needed. Other feedback from testers was better guidance for building a mesh and common mesh quality checks based on the requirements for simulation and skill testing. Dr. Joseph Zhang, SCHISM developer, provided the testers a tutorial (Zhang, 2023ab). Testers were also encouraged to attend the NOAA annual workshops for SCHISM (July 19-20, 2023) and FVCOM (August 8-9, 2023).

Challenge 2- Issues with the DEMs

Early on in the mesh generation process testers had highlighted issues with the provided DEM and bathymetry, especially along the Arthur Kill and Bayonne channels. An updated DEM was provided to the developers and the testers. This resulted in an immediate improvement in the skill assessment results, reinforcing the importance of DEM quality early on in the process. Between the eight tester groups, there were five different meshes, with three groups choosing to use the blessed mesh due to software limitations with SMS or other constraints. The tester meshes were primarily generated with SMS or OceanMesh2D, and it was highlighted by some testers that SMS has bathymetry interpolation issues for large datasets. Those with limitations preventing them from obtaining SMS have had more trouble with mesh generation than those with access. Figures 3 and 4 display the two blessed meshes for SCHISM and FVCOM, respectively.

4.3. Lessons Learned

Based on the success and challenges mentioned above, several lessons learned will be implemented in the next rounds. The three categories of lessons learned are: meetings and structure, mesh development and DEM/bathymetry suggestions, and pre/post-processing techniques.

Lesson 1- Meetings

The success of the Monthly meetings has proved itself as a good communication tool. We will continue having bi-weekly meetings hosted by a fellow student to prompt a lower-stakes environment for collaboration and questions. This also allows students to get experience in public speaking and project management. With both of these opportunities for feedback and collaboration, more successful progress has been made and new testers are able to get a developing understanding of the model they chose to implement.

Lesson 2- Clearer guidance and consistency

With the benefits of having freedom and flexibility of the testers in their configurations and methodology, there is a need for consistency. Thus, it might be a good idea for the UFS CAT to provide more support, guidance, and feedback to the testers. The DEM case is a good example for providing more guidance for mesh generation and addressing any issues with the data. It was also good that the developers jumped in to support the testers.

5. NEXT STEPS AND RECOMMENDATIONS

The model evaluation is already in the second round that will incorporate the atmospheric forcing components onto the existing tidal model, and switch to running the model in three dimensions (3D) (*i.e.*, full circulation model). The co-leads will provide atmospheric (GFS, HRRR), ocean (Global RTOFS), and river (National Water Model, USGS) forcing data as well as additional observational data (water levels, water currents, water temperature, salinity) for the testing. Test results will be provided in a future report.

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** and **Yunfang Sun**, 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.

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.

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.

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.

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

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.

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

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

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

U.S. Coast Guard Search and Rescue (2022). CG-SAR-CF-20220208 and Dr. Cristina Forbes, personal communication, June 24, 2021.

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

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