2021

Real-time environmental forecasts of the Chesapeake Bay: Model setup, improvements, and online visualization

Aaron Bever
Marjorie A.M. Friedrichs
Virginia Institute of Marine Science
Pierre St-Laurent
Virginia Institute of Marine Science

Follow this and additional works at: https://scholarworks.wm.edu/vimsarticles

Part of the Environmental Monitoring Commons

Recommended Citation
Bever, Aaron; Friedrichs, Marjorie A.M.; and St-Laurent, Pierre, Real-time environmental forecasts of the Chesapeake Bay: Model setup, improvements, and online visualization (2021). Environmental Modelling & Software, 140, 105036.

This Article is brought to you for free and open access by the Virginia Institute of Marine Science at W&M ScholarWorks. It has been accepted for inclusion in VIMS Articles by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.
Real-time environmental forecasts of the Chesapeake Bay: Model setup, improvements, and online visualization

Aaron J. Bever\textsuperscript{a,}\textsuperscript{*}, Marjorie A.M. Friedrichs\textsuperscript{b}, Pierre St-Laurent\textsuperscript{b}

\textsuperscript{a} Anchor QEA, LLC, 130 Battery Street, Suite 400, San Francisco, CA 94111, USA
\textsuperscript{b} Virginia Institute of Marine Science, William & Mary, Gloucester Point, VA, USA

\textbf{A R T I C L E   I N F O}

\textbf{Keywords:} Modeling, Forecast, Chesapeake Bay, Hypoxia, Dead zone, Acidification

\textbf{A B S T R A C T}

Daily real-time nowcasts (current conditions) and 2-day forecasts of environmental conditions in the Chesapeake Bay have been continuously available for 4 years. The forecasts use a 3-D hydrodynamic-biogeochemical model with 1–2 km resolution and 3-D output every 6 h that includes salinity, water temperature, pH, aragonite saturation state, alkalinity, dissolved oxygen, and hypoxic volume. Visualizations of the forecasts are available through a local institutional website (www.vims.edu/hypoxia) and the MARACOOS Oceans Map portal (https://oceansmap.maracoos.org/chesapeake-bay/). Modifications to real-time graphics on the local website are routinely made based on stakeholder input and are formatted for use on a mobile device. Continuous model input files were developed from daily real-time forecast input files, for hindcast simulations and efficient evaluation and improvement of the real-time model. This manuscript describes the setup of the environmental forecasting system, how the model accuracy has been improved, and the revision of online graphics based on stakeholder feedback.

1. Introduction and motivation

The Chesapeake Bay (Bay, Fig. 1) is the largest, most productive, and most biologically diverse estuary in the continental United States, providing crucial habitat and natural resources for native and migratory species (Boesch et al., 2001; Kemp et al., 2005). Natural economic benefits derived from the Bay are estimated to be valued at more than $100 billion annually (CBF, 2014). The Bay supports economically important fisheries, with blue crabs, Striped Bass and oysters generating the greatest revenue (Dewar et al., 2009). Shellfish aquaculture is also growing rapidly (Hudson and Murray, 2016). In addition, Bay waters enhance coastal property values and support a vital tourist economy, including nature-based recreation industries (Klemick et al., 2018).

The many uses of the Bay result in diverse groups of stakeholders interested in both protecting and using the Bay. However, temporally and spatially varying environmental conditions can impact how different stakeholders make daily decisions regarding their usage of the Bay’s resources. For example, seasonal hypoxia (dissolved oxygen (DO) < 2 mg/L) occurs annually between May and October in the deeper channel of the Bay (Hagy et al., 2004; Officer et al., 1984). Even in the absence of direct mortality, hypoxia reduces the catch per unit effort of bottom-feeding fish and constrains the locations of productive fishing grounds (Buchheister et al., 2013). Anglers are seeking ways to receive improved information on temperature, salinity, and DO so they can make informed decisions in near real time. Additionally, both native and cultured oysters are vulnerable to negative effects from coastal acidification (Beck et al., 2009; Barton et al., 2015). A link between episodes of poor water quality and enhanced oyster mortality has been noted in recent years, although the problem is not yet well understood (Wheeler 2011; Munroe 2013). Hatchery operators have expressed enthusiasm for web-based forecasts of key environmental parameters to help guide their daily decision making.

Real-time environmental forecasting systems based on numerical hydrodynamic and biogeochemical (water quality) models have the potential to provide valuable information to both assist stakeholders in planning their daily activities and help decision-makers continuously track environmental conditions in real time. For example, LiveOcean forecasts environmental conditions throughout Puget Sound, the coastal ocean, and Willapa Bay (LiveOcean, 2020). Another example is the Chesapeake Bay Operational Forecast System (CBOFS) which is run every 6 h by the National Oceanic and Atmospheric Administration (NOAA) and is used primarily for forecasting water levels for navigation.

\textsuperscript{*} Corresponding author.
\textit{E-mail address:} abever@anchorqea.com (A.J. Bever).

https://doi.org/10.1016/j.envsoft.2021.105036

Accepted 16 March 2021

Available online 20 March 2021

1364-8152/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
but is also used to evaluate the chance of encountering stinging jellyfish (NOAA, 2020a; NOAA, 2020b). Other examples of real-time forecasting are the Ocean Circulation, Ecosystem and Hypoxia around Hong Kong waters system (HKUST, 2020) and various systems in the Mediterranean Sea (Tintore et al., 2019; MFS, 2020). Forecasts of the Mediterranean Sea are used by a variety of stakeholders, including port managers for evaluating extreme events, tourism operators and recreational users to understand conditions on the water, and various users for responding to emergencies (e.g., oil spills) (Tintore et al., 2019).

This paper describes a real-time Chesapeake Bay Environmental Forecast System (CBEFS) and discusses recent improvements to the forecast system, both in model accuracy and graphical visualizations for stakeholders. CBEFS has provided daily nowcasts (current conditions) and 2-day forecasts of environmental conditions throughout the Bay since February 2017, with model output and online graphics formatted based on stakeholder requests and continually expanded through time. CBEFS uses a 1–2 km horizontal grid and provides 6-hourly 3-D output. The starting CBEFS configuration was based on the well-established Chesapeake Bay Regional Ocean Modeling System (ChesROMS) Estuarine Carbon Biogeochemistry (ECB) model (Feng et al., 2015). However, CBEFS uses different inputs than the published ChesROMS-ECB setups used in other applications (e.g. Da et al., 2018); therefore, it has been independently evaluated to ensure similar model accuracy to other implementations, as well as to other well-established models of the Bay (Irby et al., 2016) that only run hindcasts (simulations of past conditions).

2. CBEFS implementation and configuration

2.1. Hydrodynamic and biogeochemical model implementations

CBEFS is based on a 3-D Chesapeake Bay implementation of the open-source community Regional Ocean Modeling System (ChesROMS; Xu et al., 2012) hydrodynamic model, including 20 vertical levels on a horizontal grid with highest resolution (430 m) in the northern Bay and roughly 1 km resolution in the middle and southern Bay (Fig. 2). The main ROMS input file with the input parameters is provided in the supplementary information. CBEFS uses the ECB and Simplistic Respiration Rate (SRM) modules to simulate biogeochemistry (with DO) and only DO, respectively. The average DO from the ECB and SRM modules is used for graphics and hypoxic volume calculations.

ECB is a full biogeochemical module that contains inorganic and organic carbon and nitrogen state variables, including particulate (detritus, phytoplankton, and zooplankton) and dissolved forms (nitrate, ammonium, dissolved inorganic carbon, and dissolved organic matter) (Feng et al., 2015; Irby et al., 2016; Da et al., 2018). In addition, inorganic suspended solids, DO, and alkalinity are included as state variables. The inclusion of inorganic carbon and alkalinity as state variables is critical to successfully simulating the CO$_2$ system and associated acidification metrics (pH, aragonite saturation state ($\Omega_{AR}$)). ECB has continually been improved as part of ongoing research (St-Laurent et al., 2020; Moriarty et al., 2021; Kim et al., 2020; Turner et al., 2021). Following relevant modifications to ECB, the ECB computer code, parameters, or nutrient inputs in CBEFS are updated to the current research version to incorporate incremental improvements. Modifications to the complex ECB computer code are incorporated into the forecast system by creating a new executable based on the best-available research version and using that executable in CBEFS from that day forward. Depending on the amount of changes to ECB, updating the ROMS hydrodynamic model code may also be incorporated into the new executable.

The SRM module simulates DO using a simplistic approach based on Scully (2013). As part of a multiple model intercomparison (Luettich et al., 2013, 2017), the SRM approach has been shown to simulate DO and hypoxic volumes with similar accuracy to mechanistic coupled hydrodynamic-biogeochemical models (Bever et al., 2013; Irby et al., 2016). The initial SRM approach based on Scully (2013) used a temporally- and spatially-constant respiration rate. However, the respiration rate prescribed to simulate summer-time hypoxia was too high for simulating winter bottom DO. In the CBEFS implementation of the SRM module, DO is consumed using a prescribed spatially uniform but temporally-varying respiration rate that repeats each year. The prescribed respiration rate is higher in summer than in winter, which better captures the seasonal patterns in bottom DO. DO in the model surface layer is set to saturation based on water temperature and salinity.

The CBEFS standalone hydrodynamic and biogeochemical ROMS modeling system requires nowcast and forecast boundary conditions (input files) daily (Table 1). A technique was developed to utilize the hydrodynamic input files already being created by NOAA for CBOFS to use as boundary conditions for CBEFS, thus leveraging work already being conducted operationally by NOAA. CBOFS runs every 6 h and simulates hydrodynamic conditions throughout the Bay (Lanerolle et al., 2009, 2011; NOAA, 2020a). The input files for CBOFS are retrieved from NOAA (2020c), appended together, and reformatted using a combination of shell scripts, Network Common Data Form (NetCDF) Operators, and MATLAB scripts. These reformatted files provide the inputs for the hydrodynamic component of CBEFS. Meteorology is from the North American Mesoscale (NAM) model (EMC, 2020). Tributary freshwater inflow and temperature are based on observed U.S. Geological Survey

![Fig. 1. Chesapeake Bay and relevant tributaries.](image)
Environmental Modelling and Software 140 (2021) 105036

Environmental Modelling and Software 140 (2021) 105036

(USGS) gauge data (USGS, 2020), with the inflow volumes then scaled to better capture the total terrestrial freshwater inflow to the Bay (described in Section 3.1.1). Freshwater inflows for the forecast period of each daily simulation are held constant, based on the inflows during the nowcast period. Tides at the open boundary are derived from the Advanced Circulation model (ADCIRC, Luettich et al., 1992) and non-tidal water levels are from the Extratropical Storm Surge Model (OPC, 2019). Ocean boundary temperature and salinity are from the Global operational Real-Time Ocean Forecast System (RTOFS) (NWS, 2019).

The inputs necessary for the CBEFS implementation of ECB are the same as those necessary for the research version referenced above. However, because CBOFS includes only hydrodynamic fields, biogeochemical inputs for CBEFS must be obtained from other sources. Riverine concentrations of biogeochemical variables (excluding DIC and alkalinity) for the 13 tributaries included in CBEFS are specified based on climatological values derived from the Dynamic Land Ecosystem Model (DLEM; Tian et al., 2015; Yang et al., 2015a, 2015b; Feng et al., 2015). DIC and alkalinity riverine inputs are based on St-Laurent et al. (2020) and repeat 2014 for each subsequent year. At the ocean boundary, climatological values based on Da et al. (2018) and St-Laurent et al. (2020) are used. Tributary and ocean boundary DO concentrations are set to saturation. Atmospheric nitrogen deposition is based on Da et al. (2018) and repeats the last available year (2014) for each subsequent year. Input parameters for the ECB model are the same as described in St-Laurent (2020). The ROMS biology.in file with parameter values is provided in the supplemental information.

2.2. Forecast system configuration

CBOFS runs in a Linux environment and uses the cron software utility to automatically run shell scripts at specific times of the day to completely automate the CBEFS workflow (Fig. 3). NetCDF operators (i.e., NetCDF kitchen sink (ncks)) are used to efficiently modify NetCDF input and output files. The sed command is used to replace dates in a generic ROMS text input file so the simulation starts on the correct day. MATLAB is used to further preprocess/postprocess model input/output files and generate portable network graphics (png) files for online presentation.

![Fig. 2. General setup of CBEFS.](chart)

<table>
<thead>
<tr>
<th>Input</th>
<th>Variables</th>
<th>Initial Source</th>
<th>Source</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlantic Ocean Boundary</td>
<td>Tides</td>
<td>Advanced Circulation Model</td>
<td>37 Tidal Harmonics; Variable (~5 km near Chesapeake Bay Open Boundary)</td>
<td>Hourly; Interpolation Between Duck NC and Ocean City MD</td>
</tr>
<tr>
<td></td>
<td>Non-tidal Water Levels</td>
<td>Extratropical Storm Surge Model</td>
<td>Hourly;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Salinity and Temperature</td>
<td>Global operational Real-Time Ocean Forecast System</td>
<td>Hourly;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Biogeochemistry</td>
<td>Climatology</td>
<td>1/12</td>
<td></td>
</tr>
<tr>
<td>Atmospheric</td>
<td>Meteorology</td>
<td>North American Mesoscale Model</td>
<td>3-hourly; 12 km (Prior to January 2018)</td>
<td>Hourly; 3 km (After January 2018) Daily; Spatially Constant</td>
</tr>
<tr>
<td></td>
<td>Nitrogen Deposition</td>
<td>Climatology from Da et al. (2018)</td>
<td>Hourly</td>
<td>Daily; Spatially Constant</td>
</tr>
<tr>
<td>River Inflow</td>
<td>Discharge</td>
<td>USGS</td>
<td>Hourly</td>
<td>Daily or Hourly; Spatially Variable</td>
</tr>
<tr>
<td></td>
<td>Temperature</td>
<td>Climatology or USGS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Biogeochemistry</td>
<td>Dynamic Land Ecosystem Model Climatology</td>
<td>Temporally Constant to Monthly; Spatially Constant or Spatially Variable</td>
<td></td>
</tr>
</tbody>
</table>
visualization. The CBEFS workflow uses Linux command line function calls whenever possible because they can be completed without having to request high-performance computing (HPC) time through a job scheduler or use a separate dedicated workstation computer (i.e., minimize the use of MATLAB), and thus can be done more efficiently and reliably.

Using local HPC resources, CBEFS simulates 3 days nightly, including a 1-day nowcast and 2-day forecast. The nowcast is a simulation extending through midnight using the best available inputs. Each successive nowcast restarts from the end of the nowcast for the prior day. Throughout the day, the forecast system retrieves necessary information from a NOAA ftp page for model inputs using the wget command. At night, the information is reformatted into CBEFS input files, the ECB and SRM simulations are run through the high-performance computing job scheduler, additional post-processing of model output is conducted, checks are conducted to determine whether the forecast was successful (with appropriate notification emails sent automatically via the shell scripts), and graphics for online visualization are generated (Fig. 3).

Following completion of the daily forecast, pH and $\Omega_{AR}$ are calculated using CO2SYS (Lewis and Wallace, 1998), as implemented in MATLAB, and appended to the ECB NetCDF output files using a NetCDF operator. Graphics designed for online visualization are created from the CBEFS NetCDF output files via Linux shell scripts running MATLAB scripts. Image files of the graphics are displayed in real time on the internet through a local institutional website (VIMS, 2020) and separated into various pages based on the information conveyed. The website and graphics are designed for ease of use on a mobile device and are revised as stakeholders provide feedback on what is most useful and easily interpreted (Section 4).

The Mid-Atlantic Regional Association Coastal Ocean Observing System (MARACOOS) downloads the model output daily, with visualization of the forecast output available through the Chesapeake Bay Oceans Map webpage (MARACOOS, 2020a). This webpage allows for viewing each of the vertical levels of model output, simple real-time model-data comparison at discrete locations, and comparing temperature and salinity from CBEFS to CBOFS at discrete locations. The VIMS and MARACOOS websites are complementary and provide different features for use by stakeholders. Instantaneous NetCDF output files for select CBEFS variables are provided publicly through a Thematic Real-time Environmental Distributed Data Services (THREDDS) data server (MARACOOS, 2020b).

3. Improvement of the real-time forecast system: model accuracy and refinement of model inputs

3.1. Model improvements

Accurate nowcasts (current conditions) are critical to accurately forecasting environmental conditions in a real-time model. Continuous input files spanning 2014 through 2017 were generated from the daily nowcast input files and were used to evaluate the forecast system and test potential improvements. This allowed for efficient continuous hindcasts to be conducted as if the simulation was being run in the forecast system. To improve model accuracy, two potential improvements were tested: 1) scaling river inflow to better represent total terrestrial freshwater inflow to the Bay; and 2) scaling wind speed to better match observed wind speeds over the Bay. Both scalings were developed to be relatively simple, facilitating incorporation into CBEFS. It is critical that these do not depend on additional real-time data or data and model results co-occurring in time because data are not available for the future to adjust forecast inputs. In sections 3.1.1 and 3.1.2 these two model improvements are individually discussed, and in the following sections (Sections 3.2 and 3.3) the increased accuracy of these improvements is examined.

3.1.1. Terrestrial freshwater inflows

The USGS terrestrial freshwater inflow gauge data used for model inputs do not capture stream inflows to tributaries below the gauges, inflows from smaller streams, and overland flow to the Bay. It was hypothesized that better capturing the total amount of terrestrial freshwater inflow to the Bay could improve model accuracy. To this end, daily-averaged terrestrial freshwater inputs from CBOFS based on USGS gauges were compared to estimates from DLEM that include these other sources of freshwater. DLEM estimated inflows were available for
the 8 primary tributaries of the 13 inflows included in the CBOFS input files, with the remaining 5 tributaries comprising only about 1% of the tributary freshwater flow to the Bay. For this comparison, all DLEM surface runoff was added to the estuarine model at the same locations as was the freshwater inputs from the CBOFS corresponding tributary. A 2-year overlap period of DLEM estimates and USGS inputs spanning 2014 and 2015 was available for the comparison. This comparison demonstrated that using only the gauge data considerably underestimated the total terrestrial freshwater inflow to the Bay.

A least-squares best-fit linear relationship between the DLEM estimates and the USGS gauge data was developed for eight tributaries, with each relationship indicating the inflow from the gauge data could generally be increased to account for both gauged versus ungauged drainage area (Fig. 4). Scaling factors were developed by calculating the inverse of the slope of the linear relationship, and constant offsets were determined using the intercept of the linear relationship (Table 2). The scaling factors were smallest for tributaries with gauge stations near the Bay with relatively little shoreline distance versus drainage area, such as the Susquehanna and Potomac Rivers. The scaling factors were largest for tributaries with gauge stations farther from the Bay and relatively large shoreline distance versus drainage area, such as the Choptank and Nanticoke Rivers. The tributary inflows derived from the CBOFS inputs were adjusted based on the scaling factors and offsets to generate inputs for CBEFS. These scaled inflows were used to examine the effect on model accuracy (Sections 3.2 and 3.3) and then incorporated into the daily CBEFS setup.

3.1.2. Wind speed

Gridded wind products often underestimate the wind speed over the waters of Chesapeake Bay, which Scully (2013) hypothesized was from not adequately representing the wind speed over water of the Chesapeake Bay relative to the wind speed over land. We hypothesized that scaling the wind speed from NAM to more closely match observed wind speeds could improve the accuracy of the model. To this end, wind speed from the NAM gridded input was matched to 14 wind data locations around the Bay and relationships between the observed wind speed and NAM wind speed were developed for the time period from 2014 through 2017. Twelve of the relationships indicated the NAM wind speed would better match observed winds if their magnitude was increased (scaling factor greater than 1), while the other two (located near the upstream end of the tidal Potomac River and near the C&D Canal) indicated a decrease in wind speed was required (scaling factor less than 1; Fig. 5). It is interesting to note the scaling factors that are less than one are located in areas with relatively little nearby water area, suggesting the scaling factors could be at least partially influenced by differences in wind speed over water relative to over land. Scaling factors from each of the 14 relationships were interpolated spatially to each of the gridded meteorology input cells. For model stability constraints, spatially-interpolated scaling factors were required to be no larger than 1.15 when adjusting the CBEFS inputs, because the CBEFS model setup does not include wetting and drying. CBEFS input NAM wind speeds were multiplied by the spatially-varying scaling factors to examine the effect on model accuracy. Wind direction was unchanged.

3.2. Hindcast simulations for evaluating model accuracy

Hindcast simulations spanning 2014 through 2017 were conducted for four scenarios, using: 1) the initial CBEFS setup based solely on the inputs reformatted from CBOFS (Initial CBEFS Setup); 2) scaled inflows only (Scaled Inflows); 3) scaled wind speed only (Scaled Winds); and 4) scaled inflows combined with scaled winds (Scaled Inflows and Winds). Model outputs from these four scenarios were compared to observed data to evaluate the model accuracy resulting from each set of inputs. Target diagram statistical analyses were used to assess model accuracy by comparing to observed bottom salinity, bottom temperature, and bottom DO collected by the long-term Water Quality Monitoring Program (WQMP) at 13 locations in the mainstem of the Bay (Irby et al., 2016). Target diagrams were used to visualize model skill graphically and quantitatively, using the standard-deviation-normalized bias ($bias_{N}$) and unbiased root-mean-squared-difference ($ubRMSD_{b}$) (Jolliff et al., 2009; Hofmann et al., 2008). The $ubRMSD_{b}$ is the RMSD after the bias between the modeled and observed values has been removed from the modeled values. The normalized RMSD ($RMSD_{N}$) mathematically represents the magnitude of the vector addition of the $bias_{N}$ and $ubRMSD_{b}$, and is depicted graphically as the distance to the center of the circle (target). Thus, model estimates falling closer to the center of the circle are more accurate, and any point falling inside the circle of radius one ($RMSD_{N}$<1) performs better than simply estimating the mean of the

---

**Table 2**

Scaling factors and offsets used to adjust USGS gauge data to better capture total terrestrial freshwater inflow. The Pamunkey and Mattaponi Rivers are summed to a single York River inflow.

<table>
<thead>
<tr>
<th>Tributary</th>
<th>Scaling Factor</th>
<th>Offset (m$^3$/s)</th>
<th>2014 to 2017 Average Flow (m$^3$/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potomac River</td>
<td>1.32</td>
<td>78.8</td>
<td>470</td>
</tr>
<tr>
<td>Susquehanna River</td>
<td>1.38</td>
<td>-130.6</td>
<td>1155</td>
</tr>
<tr>
<td>James River</td>
<td>1.76</td>
<td>-13.0</td>
<td>404</td>
</tr>
<tr>
<td>Rappahannock River</td>
<td>2.23</td>
<td>4.5</td>
<td>112</td>
</tr>
<tr>
<td>York River</td>
<td>2.27</td>
<td>1.9</td>
<td>83</td>
</tr>
<tr>
<td>Patuxent River</td>
<td>6.79</td>
<td>0.2</td>
<td>102</td>
</tr>
<tr>
<td>Choptank River</td>
<td>11.75</td>
<td>-0.1</td>
<td>83</td>
</tr>
<tr>
<td>Nanticoke River</td>
<td>33.22</td>
<td>-0.7</td>
<td>79</td>
</tr>
<tr>
<td>Total Flow</td>
<td></td>
<td></td>
<td>2488</td>
</tr>
</tbody>
</table>

---

**Fig. 4.** Comparison of DLEM and USGS flows for the Potomac River.
was evaluated for each of the 4 years individually, and for all 4 years observations (Jolliff et al., 2009; Hofmann et al., 2008). Model accuracy combined. The bias and hydrodynamic inputs and evaluate model accuracy. drodynamics that result from the specified inputs; therefore, it was not The respiration rate in the SRM module is calibrated based on the hy evaluating the effects of scaling inflows and wind speed on bottom DO. Only DO from the ECB module was used for accurately model than the surface values and stakeholders are most interested in the bottom DO. Only DO from the ECB module was used for evaluating the effects of scaling inflows and wind speed on bottom DO. The respiration rate in the SRM module is calibrated based on the hydrodynamics that result from the specified inputs; therefore, it was not appropriate to hold the SRM respiration rate constant and yet vary the hydrodynamic inputs and evaluate model accuracy.

3.3. Model accuracy and incorporation into CBEFS

On average, the Initial CBEFS Setup was biased slightly high for bottom salinity and bottom DO, with minimal bias in bottom temperature (Table 3; Figs. 6 and 7A). RMSE$_N$ values for the Initial CBEFS Setup were similar to those provided in Irby et al. (2016) for nine different models (although Irby et al., 2016, evaluated different years), suggesting the Initial CBEFS Setup had similar accuracy to other models of the Chesapeake Bay.

Although scaling the inflows (generally increasing terrestrial freshwater inflow; Scaled Inflows scenario) had little effect on modeled bottom water temperature (Table 3; Fig. 6A), it did reduce the model bias and RMSE$_N$ for bottom salinity; however, the correlation between modeled and observed salinity remained relatively unchanged, relative to the Initial CBEFS Setup (Table 3; Figs. 6B and 7). In terms of bottom DO, scaling the inflows resulted in slightly lower concentrations, together with a lower bias and RMSE$_N$ relative to the initial CBEFS setup (Table 3; Fig. 6C). Overall, the RMSE$_N$ for bottom salinity and bottom DO was the lowest in the Scaled Inflows scenario.

Scaling the wind speed (Scaled Winds scenario) also had little effect on modeled bottom water temperature (Table 3; Fig. 6A) and reduced the model bias and RMSE$_N$ for bottom salinity, with the correlation between modeled and observed values relatively unchanged, relative to the Initial CBEFS Setup (Table 3; Fig. 6B). However, scaling the wind speed resulted in increased bottom DO, increased bias, and increased RMSE$_N$ for bottom DO, relative to the Initial CBEFS Setup (Table 3; Fig. 6C).

The combination of scaling inflows and wind speed (Scaled Inflows and Winds scenario) resulted in the lowest bias in modeled bottom temperature, yet had little effect on the correlation between modeled and observed values and on the RMSE$_N$ (Table 3; Fig. 6A). The combination of scaling inflows and wind speed resulted in reduced bias and RMSE$_N$ in both bottom salinity and bottom DO, relative to the Initial CBEFS Setup (Table 3; Figs. 6B and C).

The accuracy of the model for simulating nowcast conditions was improved through examining effects of scaling river freshwater inflow and wind speed on bottom salinity, bottom temperature, and bottom DO. Scaling the inflows alone resulted in the most accurate setup for bottom salinity and bottom DO, based on RMSE$_N$ (Table 3). The decrease in bottom DO bias of 0.71 mg/L is likely an ecologically significant improvement in the modeled DO. Management goals and scientific studies of the Bay are separated by about 1 mg/L DO based on the ecology of the Bay. For example, anoxia is classified as 0–0.2 mg/L, 1 mg/L is the management threshold for deep channels during summer, 2 mg/L is commonly used to designate the upper DO limit for hypoxic waters, 3 mg/L is the management threshold for deep-water seasonal fish and shellfish use, etc. (USEPA, 2017). An improvement of model bias of around 1 mg/L is then important both ecologically and in terms of Bay management goals. An improvement in the salinity bias of 1.45 is likely an ecologically significant improvement in the salinity because it may either constrain or increase the area of suitable habitat for fishes (depending on the fish species) and can affect the water chemistry for shellfish. An improvement of 0.17 °C in the temperature bias is likely not

<table>
<thead>
<tr>
<th>Variable</th>
<th>Scenario</th>
<th>$r^2$</th>
<th>bias</th>
<th>RMSE</th>
<th>RMSE$_N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottom Temperature (°C)</td>
<td>Initial CBEFS Setup</td>
<td>0.99</td>
<td>0.30</td>
<td>0.99</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>Scaled Inflows</td>
<td>0.99</td>
<td>0.13</td>
<td>1.03</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>Scaled Winds</td>
<td>0.99</td>
<td>0.21</td>
<td>1.10</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>Scaled Inflows and Winds</td>
<td>0.99</td>
<td>0.04</td>
<td>1.07</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>Initial CBEFS Setup</td>
<td>0.75</td>
<td>2.25</td>
<td>2.85</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>Scaled Inflows</td>
<td>0.76</td>
<td>0.80</td>
<td>2.04</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>Scaled Winds</td>
<td>0.74</td>
<td>1.03</td>
<td>2.24</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>Scaled Inflows and Winds</td>
<td>0.75</td>
<td>−0.54</td>
<td>2.27</td>
<td>0.69</td>
</tr>
<tr>
<td>Bottom Salinity</td>
<td>Initial CBEFS Setup</td>
<td>0.88</td>
<td>1.04</td>
<td>1.68</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>Scaled Inflows</td>
<td>0.90</td>
<td>0.33</td>
<td>1.23</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>Scaled Winds</td>
<td>0.88</td>
<td>1.35</td>
<td>1.90</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>Scaled Inflows and Winds</td>
<td>0.88</td>
<td>0.69</td>
<td>1.46</td>
<td>0.39</td>
</tr>
<tr>
<td>Bottom Dissolved Oxygen (mg/L)</td>
<td>Initial CBEFS Setup</td>
<td>0.88</td>
<td>1.04</td>
<td>1.68</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>Scaled Inflows</td>
<td>0.90</td>
<td>0.33</td>
<td>1.23</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>Scaled Winds</td>
<td>0.88</td>
<td>1.35</td>
<td>1.90</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>Scaled Inflows and Winds</td>
<td>0.88</td>
<td>0.69</td>
<td>1.46</td>
<td>0.39</td>
</tr>
</tbody>
</table>
a notable improvement in modeled temperature but does work toward the goal of continual improvement in the forecast system and Bay modeling.

Following this analysis, the scaling of the inflows was incorporated into the real-time CBEFS setup to incorporate this improvement into the forecast system. After converting the CBOFS input for use in CBEFS, the nowcast and forecast inflows are scaled based on the scaling described in Section 3.1.1. CBEFS has been running with scaled inflows since March 2019. The comparison of modeled and observed bottom water temperature, bottom salinity, and bottom DO suggests that the CBEFS setup, both the initial setup and revised setup with scaled inflows, has similar accuracy to the range of hindcast models evaluated by Irby et al. (2016). RMSD values from the CBEFS model-data comparison were within the range of, or lower than, values presented by Irby et al. (2016). However, the evaluation of different years in this study and Irby et al. (2016) only allows for a general comparison of model accuracy.

4. Improvement and revision of graphics based on stakeholder feedback

In order to be useful for stakeholders, the information provided from CBEFS needs to be in a concise and easy to interpret format. At the time of publication of this manuscript, the primary stakeholders of the forecast system were anglers, oyster aquaculturists, Bay management, and scientists. Stakeholders were engaged through in-person focus groups and follow up emails and discussions during various stages of development of CBEFS, facilitated by outreach specialists at the Virginia Institute of Marine Science. Early feedback from these stakeholders indicated they preferred a webpage format that would be easily accessible and viewed on a mobile device, to ensure the nowcast and forecast information could be used while at a dock or on a boat. The VIMS (2020) website is formatted in this way, both in terms of the page layouts and resolution of the graphics.

Graphics are initially generated from CBEFS and discussed with stakeholders, who provide feedback for revising graphics. The initial focus of the CBEFS website was on bottom DO (Fig. 8), but feedback from anglers suggested that, based on their experience, it would be more useful to know the depth below the water surface at which oxygen concentrations exceeds 3 mg/L. With this information anglers could focus their efforts on waters with DO high enough (above 3 mg/L) for Striped Bass to be present. A minimum DO of 3 mg/L as suggested by the anglers matches well with fisheries-independent sampling that shows Striped Bass catch-per-unit-effort decreases below 3.5 mg/L (Bucheister et al., 2013). Vertical profile and map graphics were therefore developed to succinctly visualize the depth to 3 mg/L throughout the Bay (Fig. 9).

Profiles provide a detailed view in the vertical at discrete locations, while the map view allows the anglers to relocate to an area of the Bay where fish will more likely be found. These graphics are also informative for visualizing mixing due to infrequent large summer storms. Before the passing of Hurricane Isaias in August 2020, DO less than 3 mg/L extended to within 15–20 feet (4.6–6.1 m) of the water surface over a large portion of the Bay (Fig. 9A and B). The large amount of mixing resulting...
from the storm increased DO throughout the Bay. Following the passing of the storm, DO less than 3 mg/L extended to within about 30 feet (9.1 m) of the water surface and the spatial extent of the occurrence of DO less than 3 mg/L was reduced (Fig. 9C and D).

Stakeholders from the Chesapeake Bay management community also expressed interest in expanding DO visualizations to include hypoxic volume, the volume of Bay water with DO less than 2 mg/L, which can be continuously tracked by CBEFS in real time throughout the summer. Estimating and tracking hypoxic volume in real time can improve understanding of the severity of hypoxia throughout the summer because the data used to estimate hypoxic volume is collected at most once every 2 weeks and is typically not publicly released for several months after collection because of required quality control protocols. As such, it is not possible to estimate the volume of hypoxic conditions in the Bay from the observed data until after those conditions have passed. An initial webpage and graphics were created to track hypoxic volume throughout the summer and relate the amount of hypoxia to years past (Fig. 10). At the request of stakeholders and managers, graphics were

---

**Fig. 8.** Real-time graphics for bottom DO from August 7, 2020. Locations of CB5.1 and CB7.1 are the two southern most dots shown in Fig. 9, respectively. The dashed grey line indicates the day of the forecast.
then developed and added to the webpage to show model-data comparisons in near real time, comparing the model-estimated and data-estimated hypoxic volume throughout the summer (Fig. 11). To develop the data-estimated hypoxic volumes for overlay on the model-estimated hypoxic volumes, the authors are provided preliminary WQMP data shortly after it is collected and estimate a data-based hypoxic volume based on an inverse-distance weighted interpolation (Bever et al., 2013).

Focused salinity graphics are being developed as part of examining the effects of ocean acidification on shellfish and aquaculture in the Bay. Initially, pH, alkalinity, and Ω_{AR} maps were created for possible use by oyster aquaculture to understand if water conditions would be unsuitable for flow-through systems. However, as a result of several recent wet years in the region, the stakeholders also expressed interest in salinity. As a result, surface and bottom salinity maps focused on specific areas of the Bay were added to the VIMS website and preliminary time series salinity figures were developed for potential inclusion on the website.

5. Potential future additions and improvements to the forecast system

Notable areas of future expansion and improvements to the forecast system include...
system include real-time model-data comparisons to high-frequency continuous monitoring data, real-time data assimilation, increased duration of the forecasts, incorporation of fish habitat models, and a higher-resolution Bay-wide model grid or high-resolution nested grids over regions of interest to stakeholders. Real-time continuous monitoring water level, salinity, temperature, and dissolved oxygen data in the Chesapeake Bay are available through various agencies. These data could be used to evaluate model accuracy in real time, or incorporated into the nowcast portion of each daily model run to potentially improve the accuracy of the forecasts. For example, the IOOS Regional Association for the Pacific Northwest, NANOOS, has real-time model-data comparisons to mooring data at many locations and the MARACOOS.

Fig. 10. Daily hypoxic volume (top) and total annual hypoxic volume (bottom) through October 1, 2019.

Fig. 11. Daily stacked hypoxic volume (shading) for different regions of the Bay with data-based hypoxic volume estimates (dots and uncertainty bars) at the end of summer 2019. Uncertainty bars are based on the authors unpublished data.
Chesapeake Bay OceansMap webpage allows for a basic model-data comparison at select mooring locations. The 2-day duration of the forecasts could be lengthened to provide longer environmental forecasts (Ross et al., 2020). However, the lengthening of the duration of the forecasts would necessitate retrieving all the necessary model inputs from the original sources and developing methods to convert those data and model products to what is required by CBEFS. Lengthening the forecasts from 2 days to 5 days is planned but has not yet been conducted. Lengthening the duration of the forecasts would improve the ability of the forecast to be further utilized by stakeholders by providing more time to understand the upcoming water quality conditions and incorporate the information into their decision making.

Output from numerical models is more and more frequently being used in combination with fish habitat analyses and fish habitat models to estimate favorable habitat for fishes (e.g., Bever et al., 2016; Scales et al., 2017; Crear et al., 2020a, 2020b). Fish habitat models could be incorporated into CBEFS to continually estimate favorable habitat locations and track estimates of habitat area or volume through time. The probability of encountering harmful algal blooms could also be added to CBEFS to help park or beach managers and the public understand the likelihood of encountering harmful algal blooms during water-based recreation.

A higher-resolution model grid would better represent the bathymetry throughout the Bay and tributaries, which, based on Ye et al. (2018), may improve the accuracy of the model. While the current CBEFS grid is sufficient for anglers in the mainstem of the Bay, aquaculture stakeholders have requested model output more focused on their individual locations. A higher-resolution model would facilitate forecasts more focused on smaller areas than the Bay-wide or tributary-wide maps currently provided. Increasing the horizontal resolution of the model grid by about a factor of 3 in both horizontal directions is being done now but has not yet been incorporated into CBEFS.

6. Conclusions

A real-time environmental forecast system for the waters of Chesapeake Bay (CBEFS) was setup using a well-established open-source model and has been simulating environmental conditions daily since 2017. CBEFS simulates 1 day of nowcast followed by 2 days of forecast nightly, generates graphics, displays the graphics online, and makes the model output available in real time through a THREDDS server. Online graphics have continually been revised and expanded based on feedback from a diverse group of stakeholders. CBEFS leverages work already being done operationally by NOAA for the creation of daily input files for the forecast system. This technique of leveraging previous work to efficiently develop a hydrodynamic and biogeochemical forecast modeling system can be done anywhere that already has a hydrodynamic modeling forecast system, even if the hydrodynamic models are different.

Further examining the model accuracy demonstrated that scaling the tributary freshwater inflows to better match the total terrestrial freshwater inflow to the Bay improved the accuracy of the forecast system. Scaling the wind speed based on data over the Bay water improved the accuracy of bottom salinity but decreased the accuracy of bottom dissolved oxygen. Based on this analysis, the scaling of terrestrial freshwater inflow was incorporated into the forecast system to improve the accuracy. Future efforts to improve CBEFS will focus on increasing the horizontal resolution, lengthening the duration of the forecasts, and adding real-time model-data comparisons onto the website.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank David Forrest (VIMS), Raleigh Hood (UMCES), and VIMS/W&M IT for help with various aspects of the initial setup, THREDDS data server, and general IT support. David Malmquist (VIMS) has been instrumental in helping with the VIMS CBEFS website and annual Dead Zone Reports based on CBEFS. Extension Specialists Susanna Musick and Karen Hudson (VIMS) led stakeholder outreach activities for recreational anglers and aquaculturists, respectively. The authors acknowledge William & Mary Research Computing for providing computational resources and/or technical support that have contributed to the results reported within this paper (https://www.wm.edu/it/rc). This paper is the result of research funded by NOAA’s Ocean Acidification Program under award NA18OAR0170430 and NOAA’s National Center for Coastal Ocean Science under award NA16NOS4782027, both to the Virginia Institute of Marine Science. The authors also acknowledge funding from the Mid-Atlantic Regional Association Coastal Ocean Observing System under award NA16NOS0120020. This is Virginia Institute of Marine Science contribution number 4003.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envsoft.2021.105036.

Software and data availability

Documentation and source code for the numerical model used in the forecast system (ROMS) are publicly available at www.myroms.org. The cruise-based data used in this manuscript are publicly available through the Chesapeake Bay Program online data server at http://data.chesapeakebay.net/WaterQuality. Model output from the forecast system for select variables is publicly available through MARACOOS at http://data.oceansmap.com/eds_thredds/catalog/EDS/VIMS_ROMS/cat alog.html. Because of continual improvements to the forecast system, the model output on the THREDDS server may not be the product of a single consistent setup.

References
