Combining Observations & Models to Improve Estimates of Chesapeake Bay Hypoxic Volume*

Aaron Bever, Marjorie Friedrichs, Carl Friedrichs, Malcolm Scully, Lyon Lanerolle

*Submitted to JGR-Oceans
• What method(s) do you use for assessing DO and why?
• What have you found drives DO patterns in the Bay?
• What lessons have you learned?
• What method(s) do you use for assessing DO and why?
• What have you found drives DO patterns in the Bay?
• What lessons have you learned?
Objectives:

1. Use multiple models to examine uncertainties caused by interpolating hypoxic volumes, due to:
   • Data are not a “snapshot” (collected over ~2 weeks)
   • Data have coarse spatial resolution
2. Use these multiple models to correct the CBP interpolated hypoxic volumes
3. Use these corrected time series to assess different metrics for estimating interannual variability in hypoxic volume
   • Average Summer Hypoxic Volume
   • Cumulative Hypoxic Volume
Background:
The U.S. IOOS Testbed Project

Estuarine Hypoxia Team:

Marjorie Friedrichs (VIMS)
Carl Friedrichs (VIMS)
Aaron Bever (VIMS)
Jian Shen (VIMS)
Malcolm Scully (ODU)
Raleigh Hood/Wen Long (UMCES)
Ming Li (UMCES)
Kevin Sellner (CRC)

Federal partners

Carl Cerco (USACE)
David Green (NOAA-NWS)
Lyon Lanerolle (NOAA-CSDL)
Lewis Linker (EPA)
Doug Wilson (NOAA-NCBO)
Background:
The U.S. IOOS Testbed Project

Methods:

• Compare relative skill of various Bay models
• Compare strengths/weaknesses of various models
• Assess how model differences affect water quality simulations

What should a “Next Generation Bay Model” entail?
Five Hydrodynamic Models Configured for the Bay

CH3D
Cerco & Wang
USACE

UMCES-ROMS
Li & Li
UMCES

EFDC
Shen
VIMS

CBOFS (ROMS)
Lanerolle & Xu
NOAA

ChesROMS
Long & Hood
UMCES
Five Hydrodynamic Models Configured for the Bay

- CH3D
  - Cerco & Wang
  - USACE

- EFDC
  - Shen
  - VIMS

- UMCES-ROMS
  - Li & Li
  - UMCES

- ChesROMS
  - Long & Hood
  - UMCES

- CBOFS (ROMS)
  - Lanerolle & Xu
  - NOAA
Five Biological (DO) Models Configured for the Bay

- **ICM**: CBP model; complex biology
- **BGC**: NPZD-type biogeochemical model
- **1eqn**: Simple one equation respiration (includes SOD)
- **1term-DD**: depth-dependent respiration (not a function of $x$, $y$, temperature, nutrients...)
- **1term**: Constant net respiration (not a function of $x$, $y$, temperature, nutrients OR depth...)
### Coupled hydrodynamic-DO models

**Four combinations:**

- CH3D + ICM $\leftrightarrow$ CBP model
- CBOFS + 1term
- ChesROMS + 1term
- ChesROMS + 1term+DD

Physical models are similar, but grid resolution differs
Biological/DO models differ dramatically
All models (except CH3D) run using same forcing/boundary conditions, etc…
Relative Model Skill

How well do the models represent the mean and variability of dissolved oxygen at ~40 CBP stations in 2004 and 2005?

● = ~40 CBP stations used in this model-data comparison
Relative model skill: Target diagrams

Model skill (RMSD) = Distance from Origin symbol at origin → model fits observations perfectly

Jolliff et al., 2009
Relative model skill: Target diagrams

Model skill (RMSD) = Distance from Origin
symbol at origin → model fits observations perfectly

Jolliff et al., 2009

Spatial variability

CH3D + ICM

ChesROMS + Depth Dep.
CH3D-ICM and ChesROMS reproduce DO patterns similarly well
All six model combinations performed similarly well.
Objectives:

1. Use multiple models to examine uncertainties caused by interpolating hypoxic volumes, due to:
   - Data are not a “snapshot” (collected over ~2 weeks)
   - Data have coarse spatial resolution
2. Use these multiple models to correct the CBP interpolated hypoxic volumes
3. Use these corrected time series to assess different metrics for estimating interannual variability in hypoxic volume
   - Average Summer Hypoxic Volume
   - Cumulative Hypoxic Volume
Data-derived HV estimates

Data:
- Of 99 CBP stations (red dots), 30-65 are sampled each “cruise”

Note: Cruises use 2 boats from 2 institutions to collect vertical profiles; last for up to 2 weeks

Interpolation Method:
- CBP Interpolator Tool
- HV = DO < 2 mg/L
- Full Bay

Uncertainties arise from:
- Temporal errors: data are not a snapshot
- Spatial errors: discrete data cannot resolve entire Bay
Model Skill: Hypoxic Volume

Data-derived HV vs. Integrated 3D Modeled HV

However... Interpolated HV vs. Integrated HV is an apples vs. oranges comparison.
Model-derived HV estimates

Integrated 3D:
- Hypoxic volume is computed from integrating over all grid cells

Interpolated Absolute Match:
- Same 30-65 stations are “sampled” at same time/place as data are available

Interpolated Spatial Match:
- Same stations are “sampled”, but samples are taken synoptically

Interpolation Method:
- CBP Interpolator Tool
- HV = DO < 2 mg/L
- Full Bay
Model Skill Assessment for HV

Skill of Modeled Absolute Match is higher!

Absolute Match vs. Integrated 3D → uncertainties in data-derived HV
Hypoxic Volume Estimates

- Good comparison for Absolute Match
Hypoxic Volume Estimates

- When data and model are interpolated in the same way, good match.
- Interpolated HV underestimates actual HV for every cruise.
Hypoxic Volume Estimates

- When data and model are interpolated in the same way, good match.
- Interpolated HV underestimates actual HV for every cruise.
- Much of this disparity could be due to temporal errors (red bars).

![Graphs showing hypoxic volume estimates for different models and data sources.]

- Integrated 3D HV
- Absolute Match
- Spatial Match
- Spatial Match Range
- Cruise Date Range

- CH3D-ICM
- ChesROMS+1term
- Data-derived
Uncertainties in data-derived hypoxic volumes

The temporal errors from non-synoptic sampling can be as large as spatial errors (~5 km$^3$)

Spatial errors show interpolated HV is always too low (~2.5 km$^3$)
Objectives:

1. Use multiple models to examine uncertainties caused by interpolating hypoxic volumes, due to:
   • Data are not a “snapshot” (collected over ~2 weeks)
   • Data have coarse spatial resolution

2. Use these multiple models to correct the CBP interpolated hypoxic volumes

3. Use these corrected time series to assess different metrics for estimating interannual variability in hypoxic volume
   • Average Summer Hypoxic Volume
   • Cumulative Hypoxic Volume
Correcting data-derived hypoxic volumes

Reduce Temporal errors:

1. Choose subset of 13 CBP stations
2. Routinely sampled within 2.3 days of each other
3. Characterized by high DO variability
Reduce Spatial errors:

1. For each model and each cruise, derive a correction factor as a function of interpolated HV that “corrects” this data-derived HV.
Reduce Spatial errors:

1. For each model and each cruise, derive a correction factor as a function of interpolated HV that “corrects” this data-derived HV.

2. Apply correction factor to HV time-series

3. Data-corrected HV more accurately represents true HV
Interannual (1984-2012) data-corrected time series of Hypoxic Volume
Objectives:

1. Use multiple models to examine uncertainties caused by interpolating hypoxic volumes, due to:
   • Data are not a “snapshot” (collected over ~2 weeks)
   • Data have coarse spatial resolution

2. Use these multiple models to correct the CBP interpolated hypoxic volumes

3. Use this corrected time series to assess different metrics for estimating interannual variability in hypoxic volume
   • Average Summer Hypoxic Volume
   • Cumulative Hypoxic Volume
Interannual DO Assessment

- How do we determine which years are good/bad? Or whether we’re seeing a recent reduction in hypoxia?
  - Length of time waters are hypoxic
  - Percent of Bay (volume) that is hypoxic

- Choose metrics dependent on ecological function of interest:
  - Prolonged low HV could be worse for some species than an extensive short duration hypoxic event, and vice versa.

Different HV metrics can give different results for which years are “worst”
Of these three years, 1996 appears to have the least hypoxia
In 1996 Maximum HV is relatively low **BUT** Average Summer HV is relatively high; Maximum Annual HV is probably not the best DO metric.
Average Summer Hypoxia

Hypoxic volume, km$^3$

Red dashed lines denote period of “summer averaging”

2011 looks “good”, because much hypoxia occurs outside of “summer” time period.
Cumulative HV
Average Summer HV vs. Cumulative HV

- Performance of relative years changes
Average Summer HV vs. Cumulative HV

- Performance of relative years changes
- Average Summer HV doesn’t taken into account long HV duration
- If climate change affects time of onset, this will not be seen when using Avg Summer HV
Summary

- Information from multiple models (2004-2005) have been used to assess uncertainties in data-derived interpolated hypoxic volume estimates
  - Temporal uncertainties: ~5 km$^3$
  - Spatial uncertainties: ~2.5 km$^3$
  → These are significant, given maximum HV is ~10-15 km$^3$

- A method for correcting HV time series has been presented, using the model results

- Different HV metrics can give different results in terms of assessing DO improvement
  - Cumulative HV is a good way to take into account shifts in onset of hypoxia that could occur with climate change
Extra Slides
Average Summer HV vs. Cumulative HV

- Performance of relative years changes
- Average Summer HV doesn’t taken into account long HV duration
- If climate change affects time of onset, this will not be seen when using Avg Summer HV

As in previous slide, without HV correction

This demonstrates that the correction of HVs does not significantly affect the Average Summer HV vs. Cumulative HV conclusions
CBP13 scaled is now much more inline with the model estimates of 3D HV.