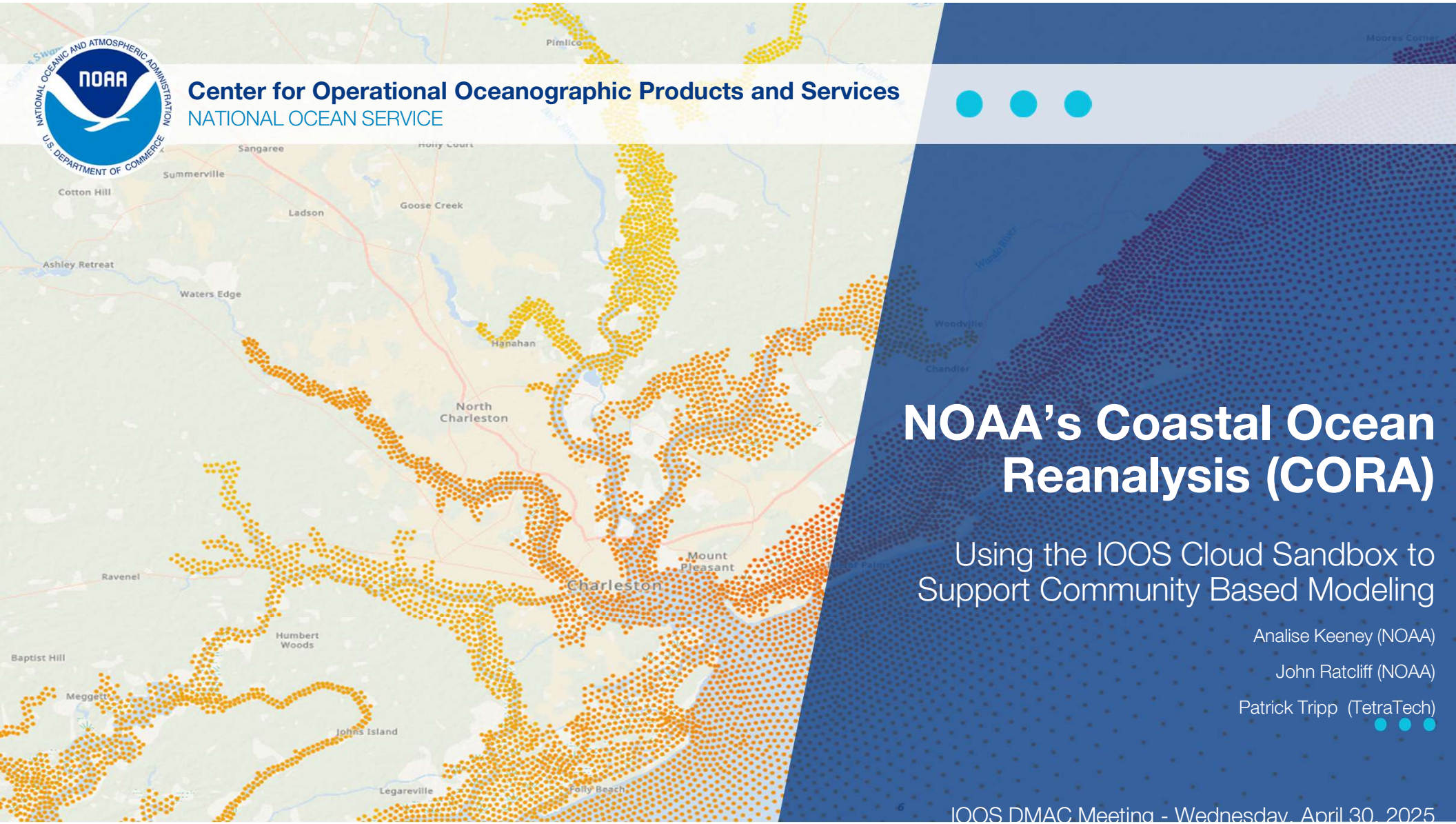


Center for Operational Oceanographic Products and Services
NATIONAL OCEAN SERVICE

NOAA's Coastal Ocean Reanalysis (CORA)

Using the IOOS Cloud Sandbox to Support Community Based Modeling

Analise Keeney (NOAA)
John Ratcliff (NOAA)
Patrick Tripp (TetraTech)

CO-OPS Impact

Grow the Blue Economy

- **Each year**, commercial mariners move **2.3 billion short tons of goods** through U.S. waterways – contributing **$397 Billion** to the U.S. GDP
- These **mariners rely daily on NOAA's network of real-time sensors** to provide **accurate navigation** data and help ensure successful deliveries

Ensure Safe Navigation

- Since its inception in 1991, **NOAA's PORTS®** Program has contributed to a **33% reduction in maritime accidents** and 59% reduction in groundings

Protect Coastal Communities

- **40%** of the **Nation's population** lives within **100 miles of the coast.**
- **CO-OPS** delivers **10+ products** designed to help these regions face their unique threats, from hurricanes and tsunamis to high tide flooding

Invest In Innovation

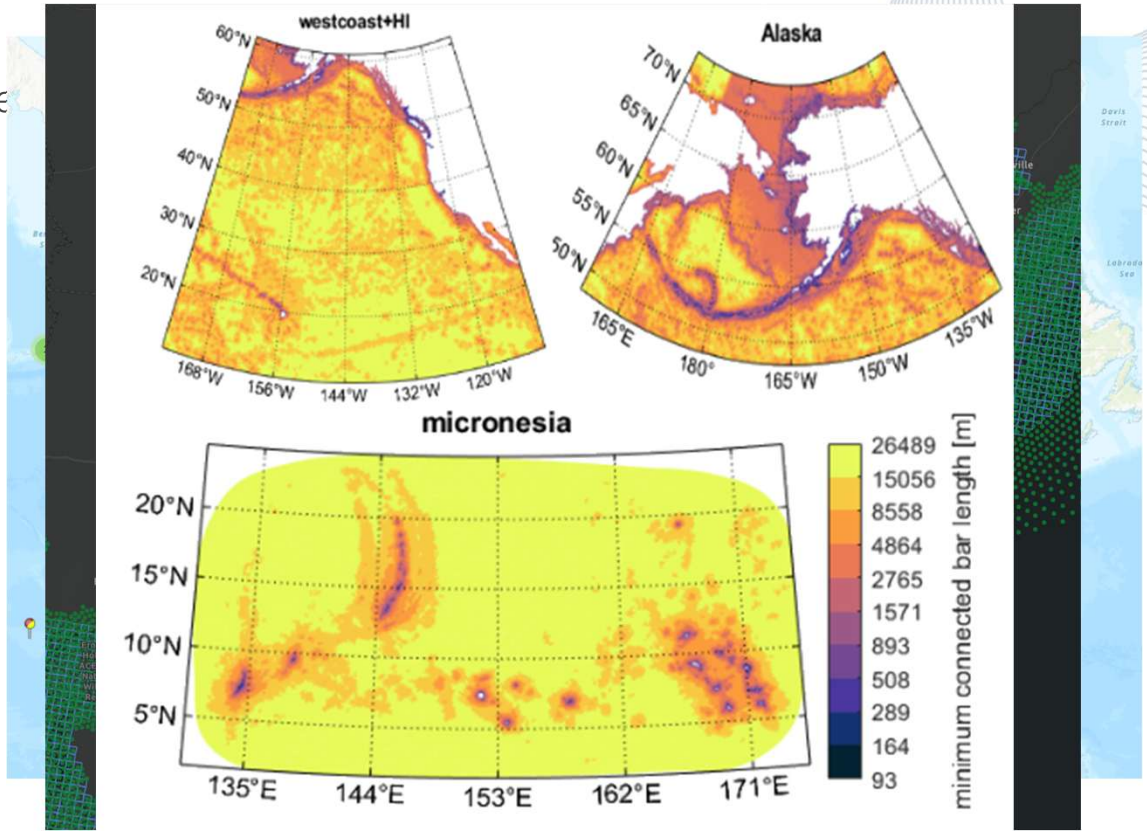
- CO-OPS is **consistently innovating** – we regularly test/install cutting-edge equipment/infrastructure and research/develop new products and models to **ensure our systems stay resilient, robust, and cost-effective**

Background

What is CORA?

NOAA's **C**oastal **O**cean **R**e**A**nalysis (CORA) is a collaborative, community-based model effort designe to:

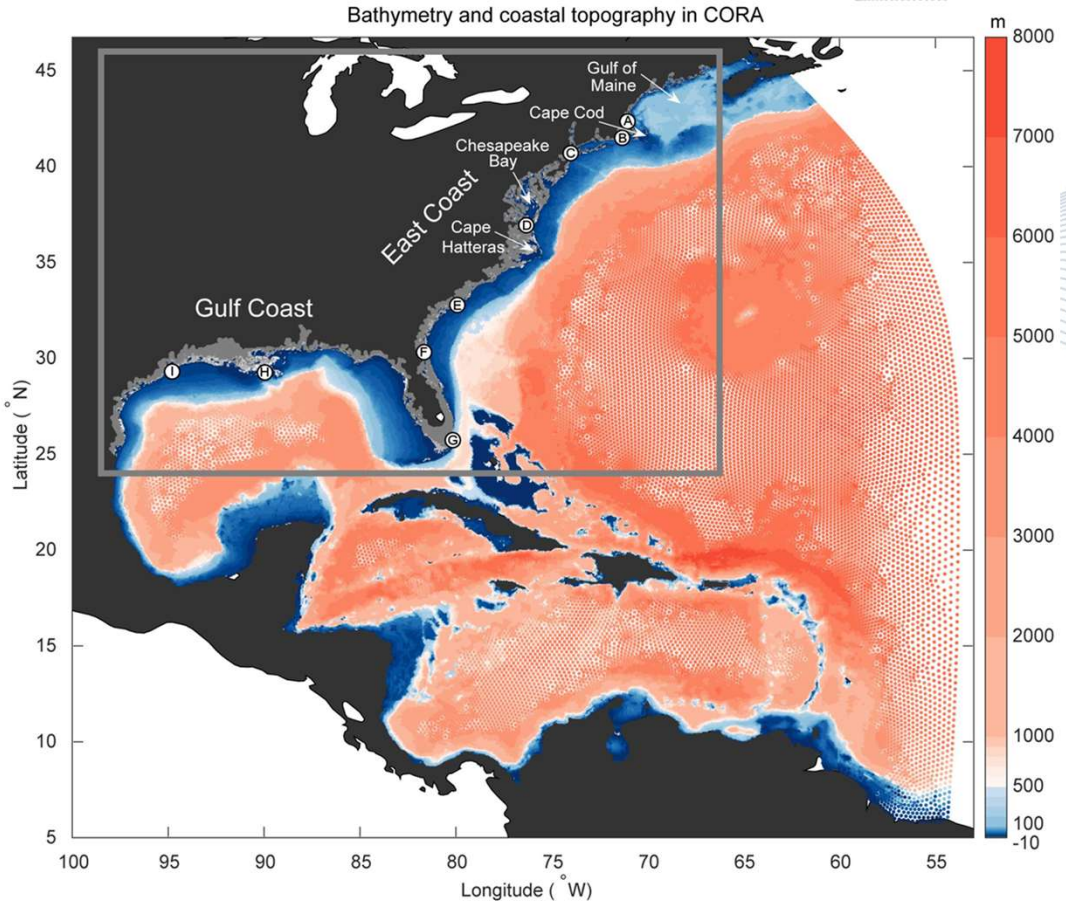
- **Leverages NWLON observations** and modern **hydrodynamic models** to **bridge gaps** in coastal long term water level observations.
- Supports improvements in **prototyping projections** and performance of **ensemble coastal water prediction** systems.
- Develops a **nation-wide dataset** that **better serves** coastal communities.



Purpose

Why invest in CORA?

- Serves to **advance** and uphold many of **NOS** [**Strategic Priorities**] associated with **Coastal Resilience and Comprehensive Services**.
- **Supports goals** outlined in the **Coastal Inundation at Climate Timescales** (CICT) White Paper, providing vital information for **flood risk assessment and planning**.
- Simulate **long-term water level information** critical for building **baseline climatologies**, **trends and monitoring** long-term sea level change, and **subseasonal to annual flood projections**.



Spatial domain and relative coverage of the ADCIRC model used for the Gulf, East Coast, and Caribbean reanalysis, CORA-GEC. Coastlines and continental shelves are noted in white to blue, and model domain is noted in red.

High Performance Computing Resources

Community Modeling for Operational Use

Hatteras at UNC's Renaissance Computing Institute (RENCI)

- Foundational methods
- Ongoing research

IOOS Cloud Sandbox

- Analysis
- Validation
- Prototyping

R&D HPC Allocation to Gaea C6

- Operational runs
- Derived product development

IOOS Sandbox Infrastructure and Workflow

Patrick Tripp, Tetra Tech - Updates since last DMAC, April 2024

- NOS Sandbox deployed end of 2023
  - Deployed behind FISMA boundary, access provided through a bastion server
  - Requires CAC to access
  - Working on Okta authentication for qualified non-CAC users
- IOOS/RPS (Tetra Tech) Sandbox deployment
  - Limited external user access for COMT model development and other funded projects
  - Setup and test models in the workflow on IOOS Sandbox, then transition to NOS Sandbox for R2O and longer analysis runs.

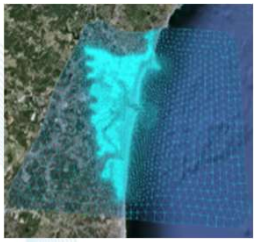
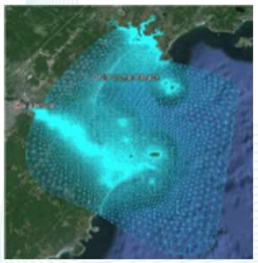
- **ADCIRC CORA in** co-operation with RENCI
- **LiveOcean Hindcasts** completed on NOS Sandbox
- From the Office of Water Prediction (**OWP**)
  - **WRF-Hydro** component of the **National Water Model**
  - **SCHISM** coastal model component of the **National Water Model**
  - **D-Flow FM** coastal model
  - Added templates for future model integrations

IOOS Sandbox Infrastructure and Workflow

- Additional AWS virtual machine instance types supported and tested
  - AMD hpc7a, Intel large memory instances, **no ARM or GPU**
  - Supported instance **options are self-discoverable** now, e.g. number of CPUs, network interfaces, etc.
  - New instance types should be **immediately usable** when released by Amazon
- Upgraded **Intel MPI**, **SPACK**, and other related libraries
- Decided to **not use SLURM**. Existing cloudflow workflow will continue on-demand cluster creation; lightweight, more flexible

- **Governance**
  - Limited permission policy to enforce **least-privilege** security, **isolated** from other AWS resources
  - **User management**, easily add multiple users, easily redeploy, using CSV file currently
- Stability and exception handling
  - Improved **"Zombie node" prevention** and management, enforce and check **exit/return values**, signal trapping
  - Automatic unique and **memorable name tags** on resources
- Improved and additional **documentation**



IOOS Sandbox Infrastructure and Workflow

Current priorities

- Model integrations in-progress:
  - **SECOFS** - South Eastern Coastal Ocean Forecast - SCHISM
  - **ECCOFS** - Eastern Coastal Community Ocean Forecast - ROMS
  - **NECOFS** - North Eastern Coastal Ocean Forecast - FVCOM

- **Training** materials and workshop
- **Contributor guides**/instructions
- **More Governance**
  - Additional **guardrails**, cost monitoring and **cost-overrun prevention**

- Re-integrate **Lustre FSx** filesystem into workflows (high-speed, low-latency, filesystem)
  - Workflow tasks will need to copy/move input and output files to/from temporary high-speed scratch disk

- **NOSOFS** Upgrade and prep for R&D
- **NOSOFS** Skill Assessment System Sandbox integration (**CO-OPS, OCS**)

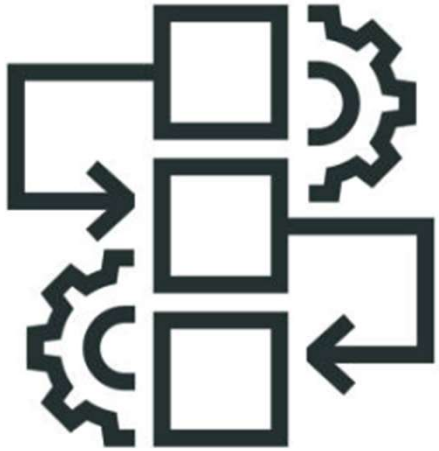
Running CORA in the Sandbox

John Ratcliff, NOAA CO-OPS

- **Model Forcing for CORA**
  - Tides, Meteorology, Waves

- **ADCIRC Input Files**
  - Fort.{13,14,15, 22, 26, 221.nc, 222.nc}, swaninit

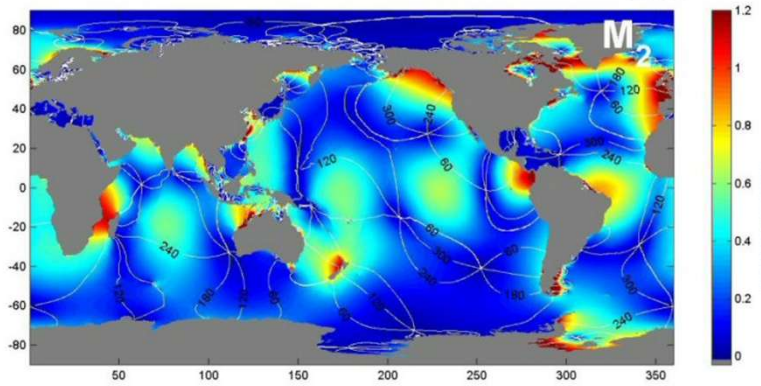
- **CORA Workflow**
  - Define compute resources
  - Define job parameters
  - Starting the simulation
  - run_storms.sh ADCIRC management script
  - File management

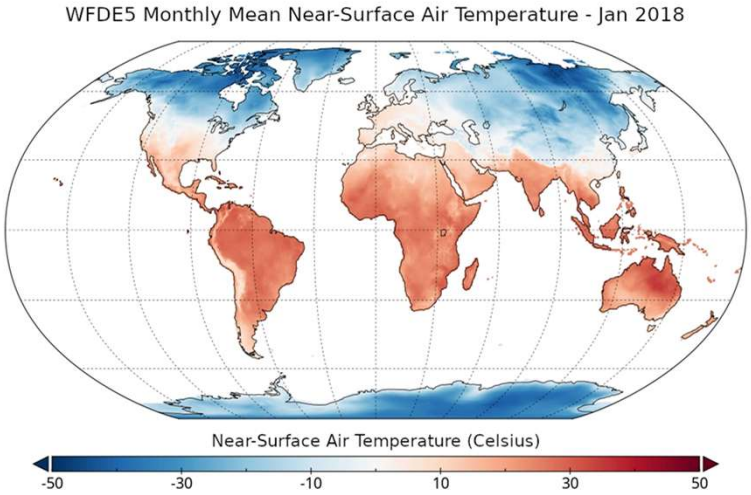


Running CORA in the Sandbox

Model Forcing

- **Tides**
  - **TPXO Global Tidal Model** for amplitude and phase at boundary nodes
  - **tide_fac.f** for nodal factors and equilibrium arguments
- **Meteorology**
  - **ERA5** Wind Velocity and Pressure Data
- **Waves**
  - **SWAN** model coupled with ADCIRC



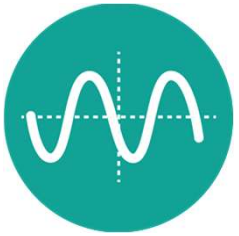


Running CORA in the Sandbox

ADCIRC Input Files

Required:

- **Fort.14**
  - Grid and Boundary Information File
  - Gulf and East Coast used widely-tested HSOFS mesh
  - Pacific is using new mesh cut from STOFS Global mesh
- **Fort.15**
  - Model Parameter and Periodic Boundary Condition File

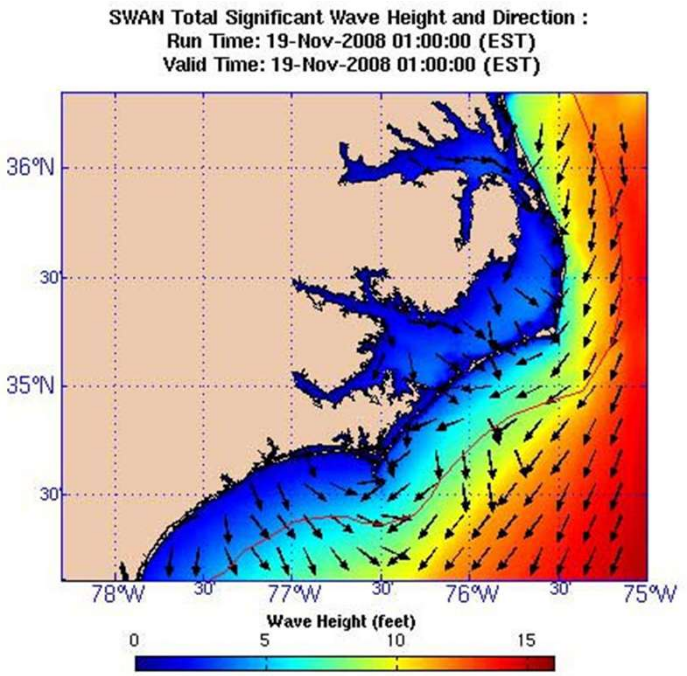


Running CORA in the Sandbox

ADCIRC Input Files

Optional/Forcing:

- **Fort.13**
  - Nodal Attributes File
  - Manning's n, surface roughness, etc.
- **Fort.26**
  - SWAN parameter file
- **swaninit**
  - SWAN init file
- **Fort.22**
  - Wind forcing file
- **Fort.221.nc and Fort.222.nc**
  - ERA5 NetCDF Pressure and Wind



Running CORA in the Sandbox

CORA Workflow

The **Sandbox** uses **Prefect** to manage workflows and tasks (**cloudflow**) in combination with the ADCIRC **run_storms.sh** script to submit and run **ADCIRC + SWAN**

Running CORA in the Sandbox

CORA Workflow: Define the Compute Resources

Cluster Config File: ec2-user/hsofs.Cloud-Sandbox/cloudflow/cluster/configs/NOS/cora.hsofs.cfg

Choose 'nodeType' and 'nodeCount'

```
"platform"   : "AWS",
"region"     : "us-east-2",
"nodeType"   : "hpc6a.48xlarge",
"nodeCount"  : 8,
"tags"       : [
                 { "Key": "Name", "Value": "CORA-node" },
                 { "Key": "Project", "Value": "NOS-cloud-sandbox" }
               ],
"image_id"   : "ami-082e2e4e7ef1c6d77",
"key_name"   : "rsa_sandbox_key",
"sg_ids"     : ["sg-05c044182398b2a27", "sg-0e5148638d9196f69", "sg-06ca6bc5d4b377dad"],
"subnet_id"  : "subnet-00075cfbfcbc8f2cf",
"placement_group" : "ioos_cloud_sandbox_Terraform_Placement_Group"
```

'hpc6a.48xlarge' instance has 96 cores and 384 GiB of memory

Running CORA in the Sandbox

CORA Workflow: Define the Job Parameters

Job File:

ec2-user/hsofs.Cloud-Sandbox/cloudflow/job/jobs/ec2-user.test.hsofs (GEC) ec2-user/hsofs.Cloud-Sandbox/cloudflow/job/jobs/test.pac-cora.reanalysis (Pacific)

```
"JOBTYPE"    : "adcircreanalysis",
"OFS"        : "adcirc-cora",
"YYYY"       : "2006",
"BUCKET"     : "",
"BCKTFLDR"   : "",
"OUTDIR"     : "",
"ProjectName" : "Reanalysis-test",
"ProjectHome" : "/save/ec2-user/cora-runs/PAC-ADCIRC",
"ADCIRCHome"  : "/save/ec2-user/adcirc/build",
"GRID"       : "pac-cora",
"CONFIG"     : "/save/ec2-user/cora-runs/PAC-ADCIRC/configs/pac-cora_prior_config.yml",
"CONFIGTMPL"  : "/save/ec2-user/hsofs.Cloud-Sandbox/cloudflow/job/templates/pac-cora_prior_config.yml.in",
"contact"    : "john.ratcliff@noaa.gov",
"WRITERCORES" : "2"
}
```

Running CORA in the Sandbox

CORA Workflow: Define the Job Parameters

Job File: ec2-user.test.hsofs (Gulf/East Coast), test.pac-cora.reanalysis (Pacific)

Choose Year to Run

```
"JOBTYPE"    : "adcircreanalysis",
"OFS"        : "adcirc-cora",
"YYYY"       : "2006",
"BUCKET"     : "",
"BCKTFLDR"   : "",
"OUTDIR"     : "",
"ProjectName" : "Reanalysis-test",
"ProjectHome" : "/save/ec2-user/cora-runs/PAC-ADCIRC",
"ADCIRCHome"  : "/save/ec2-user/adcirc/build",
"GRID"       : "pac-cora",
"CONFIG"     : "/save/ec2-user/cora-runs/PAC-ADCIRC/configs/pac-cora_prior_config.yml",
"CONFIGTMPL"  : "/save/ec2-user/hsofs.Cloud-Sandbox/cloudflow/job/templates/pac-cora_prior_config.yml.in",
"contact"    : "john.ratcliff@noaa.gov",
"WRITERCORES" : "2"
}
```

Running CORA in the Sandbox

CORA Workflow: Define the Job Parameters

Job File: ec2-user.test.hsofs (Gulf/East Coast), test.pac-cora.reanalysis (Pacific)

Choose Config Files to Use, make changes in the 'CONFIGTMPL'

```
"JOBTYPE"    : "adcircreanalysis",
"OFS"        : "adcirc-cora",
"YYYY"       : "2006",
"BUCKET"     : "",
"BCKTFLDR"   : "",
"OUTDIR"     : "",
"ProjectName" : "Reanalysis-test",
"ProjectHome" : "/save/ec2-user/cora-runs/PAC-ADCIRC",
"ADCIRCHome"  : "/save/ec2-user/adcirc/build",
"GRID"       : "pac-cora",
"CONFIG"     : "/save/ec2-user/cora-runs/PAC-ADCIRC/configs/pac-cora_prior_config.yml",
"CONFIGTMPL"  : "/save/ec2-user/hsofs.Cloud-Sandbox/cloudflow/job/templates/pac-cora_prior_config.yml.in",
"contact"    : "john.ratcliff@noaa.gov",
"WRITERCORES" : "2"
```

‘CONFIGTMPL’

```
# NTIP if true, compute tidal potentia
# expects <TidalPotentialBlock> as a t
# assumes the template fort.15 file ta
NTIP: true
# NBFR if true, expects to fill <Tidal
s template file specifies as needed.
NBFR: true
# Winds
NWS: "14"
# "hbl" | "owi" | "gahm"
WinPreSrc: "netcdf"
WithBasin: 1
WithRegion: 0
BasinPreFile: "fort.221.nc"
BasinWinFile: "fort.222.nc"
RegionPreFile: "fort.223"
RegionWinFile: "fort.224"
IncludeRivers: 1
#DRAMP: "7. 0.25 0. 0. 7. 7. 1. 0. <Rn
# forcing rampup length, days
DRAMP: 10.0
# model time step, secs
DT: 6.0
  # use padcswan (1), or not (0)
Coupled: 1
# Output
StationFile: pac-cora.stationfile.dat
#sets output SPOOLing, secs
GlobalOutputInterval: 3600.0
LocalOutputInterval: 3600.0
RNDAY: 400.0
TOUTSE: 30.9583333
DynWatLevCor: None
```

Fort.15 Filled based on ‘CONFIGTMPL’

```
<DT>                          ! DT - TIME STEP (IN SECONDS)
0.00                          ! STATIM - STARTING TIME (IN DAYS)
0.00                          ! REFTIM - REFERENCE TIME (IN DAYS)
3600.00 3600.00               ! WTIMINC
<RNDAY>                         ! RNDAY - TOTAL LENGTH OF SIMULATI
<DRAMP>                         ! DRAMP - DURATION OF RAMP FUNCTIO
0.8 0.2 0.0                   ! TIME WEIGHTING FACTORS FOR THE GWC
0.10 0 0 0.01                 ! H0, NODEDRYMIN, NODEWETMIN, VELMIN
222.286582 46.136879          ! SLAM0,SFEA0 - CENTER OF CPP PROJEC
0.00050                       ! FFACTOR
-0.2                          ! ESL - LATERAL EDDY VISCOSITY COEFF
0.0                           ! CORI - CORIOLIS PARAMETER - IGNORE
<TidalPotentialBlock>
<TidalElevationBlock>
```

Running CORA in the Sandbox

CORA Workflow: Define the Job Parameters

Job File: ec2-user.test.hsofs (Gulf/East Coast), test.pac-cora.reanalysis (Pacific)

'ProjectName', 'ProjectHome', 'ADCIRCHome', 'Grid', 'WRITERCORES' and 'contact' are pulled into the 'CONFIGTMPL'

```
"JOBTYPE"     : "adcircreanalysis",
"OFS"         : "adcirc-cora",
"YYYY"        : "2006",
"BUCKET"      : "",
"BCKTFLDR"    : "",
"OUTDIR"      : "",
"ProjectName" : "Reanalysis-test",
"ProjectHome" : "/save/ec2-user/cora-runs/PAC-ADCIRC",
"ADCIRCHome"  : "/save/ec2-user/adcirc/build",
"GRID"        : "pac-cora",
"CONFIG"      : "/save/ec2-user/cora-runs/PAC-ADCIRC/configs/pac-cora_prior_config.yml",
"CONFIGTMPL"  : "/save/ec2-user/hsofs.Cloud-Sandbox/cloudflow/job/templates/pac-cora_prior_config.yml.in",
"contact"     : "john.ratcliff@noaa.gov",
"WRITERCORES" : "2"
```

Running CORA in the Sandbox

CORA Workflow: Define the Job Parameters

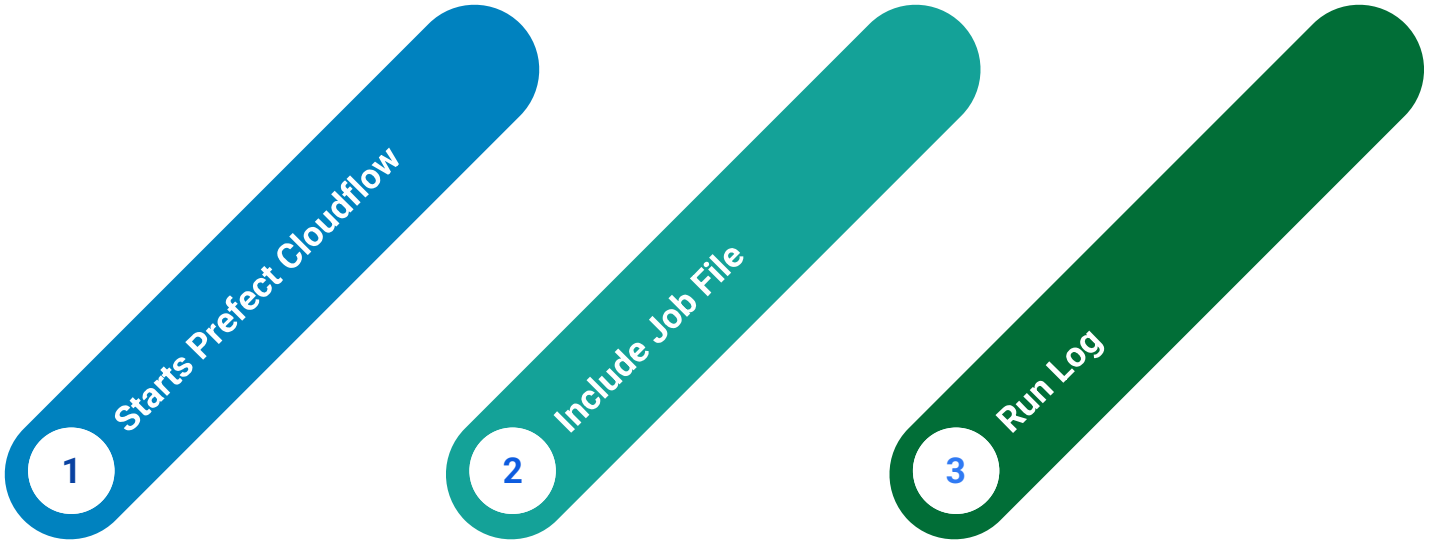
Job File: ec2-user.test.hsofs (Gulf/East Coast), test.pac-cora.reanalysis (Pacific)

'JOBTYPE' and 'OFS' are used with prefect cloudflow management to start the correct flow

```
"JOBTYPE"     : "adcircreanalysis",
"OFS"         : "adcirc-cora",
"YYYY"        : "2006",
"BUCKET"      : "",
"BCKTFLDR"    : "",
"OUTDIR"      : "",
"ProjectName" : "Reanalysis-test",
"ProjectHome" : "/save/ec2-user/cora-runs/PAC-ADCIRC",
"ADCIRCHome"  : "/save/ec2-user/adcirc/build",
"GRID"        : "pac-cora",
"CONFIG"      : "/save/ec2-user/cora-runs/PAC-ADCIRC/configs/pac-cora_prior_config.yml",
"CONFIGTMPL"  : "/save/ec2-user/hsofs.Cloud-Sandbox/cloudflow/job/templates/pac-cora_prior_config.yml.in",
"contact"     : "john.ratcliff@noaa.gov",
"WRITERCORES" : "2"
```

Running CORA in the Sandbox

CORA Workflow: Starting the Simulation



Run command: nohup ./workflows/**workflow_main.py** job/jobs/**test.pac-cora.reanalysis** >& **pacshort.out** &

Running CORA in the Sandbox

CORA Workflow: run_storms.sh

cloudflow brings in run_storms.sh

- Script system created by Brian Blanton at UNC RENCI to manage a large number of ADCIRC runs
- Uses the ‘CONFIG’/’CONFIGTMPL’ to set up ADCIRC input files and start the simulation
- Run files are written to a new directory named based on the year of the simulation

Running CORA in the Sandbox

CORA Workflow: File management

- Fort.15, Fort.22, Fort.26, and swaninit files are filled as run_storms.sh runs and output to the new run file directory

  'ec2-user/cora-runs/PAC-ADCIRC/ERA5/pac-cora/YYYY/fort.{15,22,26}'

- Fort.13 and Fort.14 are provided by user and placed in a 'Grids' directory and named based on 'GRID' in job file:

  'ec2-user/cora-runs/PAC-ADCIRC/Grids/pac-cora.13'

  'ec2-user/cora-runs/PAC-ADCIRC/Grids/pac-cora.14'

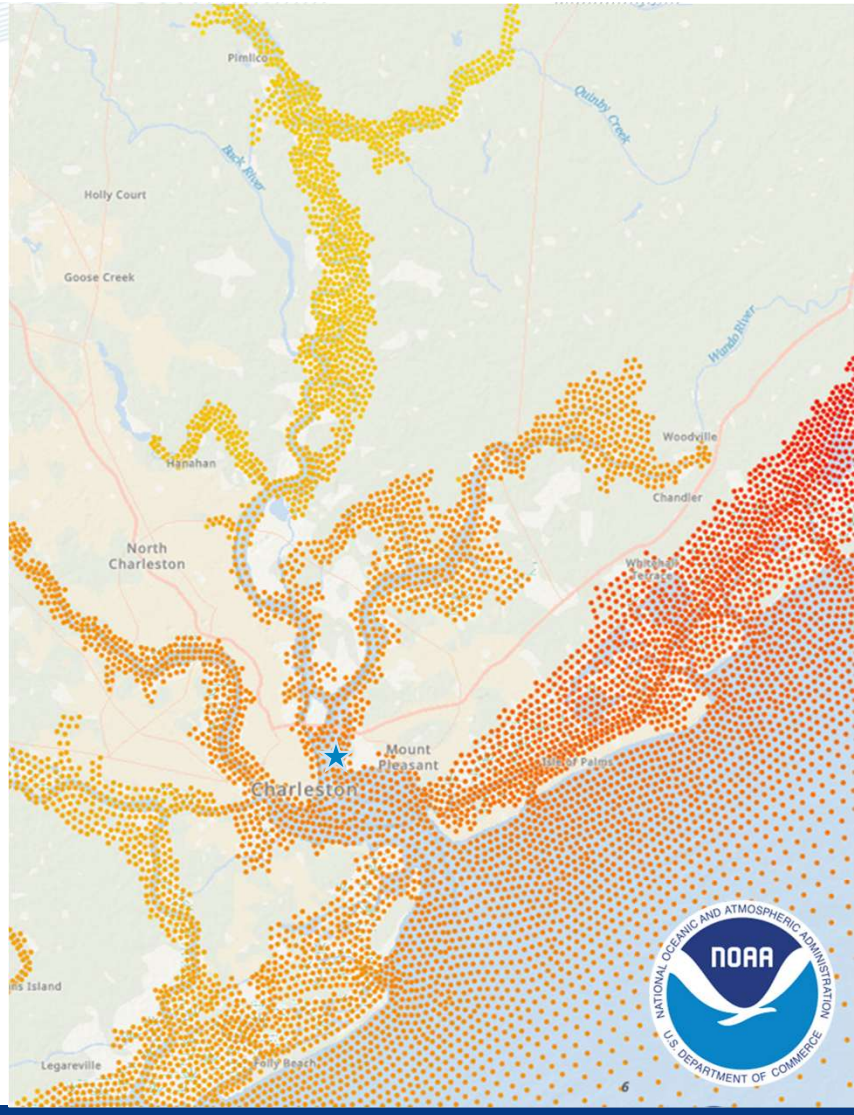
- Meteorological forcing files are provided by user placed in a 'Forcing' directory in a subdirectory named based on 'GRID' in job file: 'ec2-user/cora-runs/PAC-ADCIRC/Forcing/Winds/ERA5/pac-cora/fort.221.nc'

  'ec2-user/cora-runs/PAC-ADCIRC/Forcing/Winds/ERA5/pac-cora/fort.222.nc'

Running CORA in the Sandbox

Summary:

- The underlying hydrodynamic model for CORA is ADCIRC, which is forced by tides, meteorology, and waves (SWAN)
- Two files are required to run ADCIRC (fort.14 and fort.15) and several others are used with CORA
- CORA is run using the Cloud-Sandbox cloudflow scripts
- Primary files to set are the cluster config file, job file, and job config template file
- run_storms.sh pulls together the ADCIRC inputs and writes outputs to newly created directory for each year that is run
- Fort.{15,22,26} and swaninit files are created by run_storms, while grid and forcing files are provided by user and placed manually in appropriate directories



Community-Based Model Development

Leveraging Partnerships to Build Products

NOAA Partners:

- CO-OPS
- OCM
- IOOS
- NODD
- OAR/PSL

Academic Partners:

- University of North Carolina
- Renaissance Computing Institute
- University of Hawaii's Sea Level Center

Contract Partners:

- Tetra Tech's RPS Group

TETRA TECH

Questions?