NOAA National Weather Service (NWS) + NOAA Open Data Dissemination (NODD) + Microsoft Office Hours
October 18, 2023 | 12-1:15 PM EDT

TO: NWS NODD Microsoft Office Hours Participants

DATE: 18 OCTOBER 2023 | 12-1:15 PM EDT

FROM: Brian Cosgrove, Sudhir Shrestha, Cindy Elsenheimer, (NOAA National Weather Service), Adrienne Simonson, Jenny Dissen & Kate Szura (NOAA Open Data Dissemination Engagement and Communication)

SUBJECT: Responses to Questions from NWM Office Hours

Dear Colleagues,

Thank you again for your tremendous contribution during the NWS NODD Microsoft Office Hours. Your data related questions and comments raised during the discussion were heard and noted by NOAA.

This document provides brief responses to questions that were identified during the registration and that were raised during the discussion. Names and attributions of individuals and their affiliation have not been documented, unless it is a NOAA speaker.

We recognize the importance of continued engagement and collaboration, and invite ongoing comments via our emails.

Thank you,
Cindy, Brian, Sudhir (National Weather Service)
Adrienne, Jenny, Kate (NODD Engagement and Communication)

<u>Contents for the Summary Notes</u>

NOAA National Weather Service (NWS) + NOAA Open Data Dissemination (NODD) + Microsoft Office Hours
October 18, 2023 | 12-1:15 PM EDT

1. General Agenda of the Webinar

OUTLINE FOR THE DISCUSSION

12:00 - 12:05   Brief Introductions by NWS NWM, NODD, and Microsoft
12:05 - 12:25   National Water Model (NWM) and Update to NWM v3.0
12:25 - 12:45   NWM 3.0 Datasets Access via Microsoft (Training and Demo) and Use Cases
12:45 - 1:05   Open Discussion (Please use "Raise Hand" or the chat to raise questions)
1:05 - 1:15   Summary Comments/Closing Remarks/Next Steps

2. Questions and Responses

The questions below were identified as part of the registration process and during the Office Hours discussion. Responses are provided in brief where the NOAA team felt information was available.

QUESTIONS RAISED FROM REGISTRATION FORM

QUESTION FROM REGISTRATION	RESPONSE
I don't use the NWM personally, but am looking to find out more about how NOAA makes data available.	NWM data and related products are made available through the operational NOMADS data dissemination platform, NOAA Open Data Dissemination (NODD) cloud partners, and GIS services. The below links reflect access options and include metadata which includes details on data formats and content. For example, NOAA platforms include: https://nomads.ncep.noaa.gov/pub/data/nccf/com/nwm/prod/ https://www.weather.gov/owp/operations. https://maps.water.noaa.gov/server/rest/services/nwm Links to cloud data access points via NODD and NODD's cloud partners include: ● Azure Planetary Computer and Microsoft AI for Earth Datasets ● Google Cloud ● AWS Registry of Open Data
Storm / pSurge / Flood	Thank you for submitting this as a topic of interest. Additional information is needed to better understand the question.

Accessing data and how to work with this data in the NODD	NWM data and related products are made available through the operational NOMADS data dissemination platform, NOAA Open Data Dissemination (NODD) cloud partners, and GIS services. The below links reflect access options and include metadata which includes details on data formats and content. For example, NOAA platforms include: https://nomads.ncep.noaa.gov/pub/data/nccf/com/nwm/prod/ https://www.weather.gov/owp/operations. https://maps.water.noaa.gov/server/rest/services/nwm Links to cloud data access points via NODD and NODD's cloud partners include: ● Azure Planetary Computer and Microsoft AI for Earth Datasets ● Google Cloud ● AWS Registry of Open Data
How can this be used for emergency management operations, particularly during hurricane planning.	The NWM provides complementary, automated guidance, with the official forecast coming from NWS Weather Forecast Offices, River Forecast Centers, and National Centers like National Hurricane Center (NHC). Emergency managers should consult directly with relevant NWS offices, but can also access NWM-based products such as flood inundation maps and value-added streamflow products. With NWM v3.0, the NWM now features a total water level capability which is particularly relevant to hurricane cases wherein coastal areas experience flooding caused by the combination of storm surge, tide and freshwater river flow.
Is NOAA and NODD interested in providing users with the ability to create and / or use cloud-optimized versions of the data stored in the cloud?	NOAA National Weather Service (NWS) Office of Water Prediction (OWP) is interested in providing the data in an Analysis Ready Cloud Optimized (ARCO) data format stored in the cloud. OWP is working on chunking the NWM in a way that is cloud optimized. Please look for any updates on https://water.noaa.gov/about/nwm. NODD provides data owners and producers with feedback from users with regard to desire for Cloud Optimized formats and works with them to make these data available if they choose to convert the datasets. NOAA is actively investigating multiple avenues for providing more cloud optimized versions of their datasets based on user feedback.

	As one of NODD's cloud partners, Microsoft hosts the data in its original form, but will occasionally do cloud-optimizations steps (currently not for the NWM).
Rainfall forecasts are available (X)# of days out from an event but the water models don't reflect any farther than X # of days out. (i.e., it takes a few days for the water models to catch up with the forecast)	The length of NWM forecasts are limited by the availability of compute and forecast data. They typically run either as far out into the future as does the input forcing data coming from the weather model (same length as weather forecast) or as far as compute and time allows. If there are specific forecast configurations there are questions about, the NWM team can provide more information.
Resilience applications of the NWM	NWM long-term multi-decade retrospective simulations can be used to help assess the relative flood risk for different locations, and can put real-time analyses and forecasts into historical context. Real-time analyses can aid in disaster recovery, depicting flood-impacted regions, and real-time forecasts can help inform the pre-positioning of disaster response supplies. Additionally, with the brand new total water level capability of the NWM, the system is beginning to provide added insight into the additive impacts of storm surge and freshwater flooding for coastal communities.
Looking to learn more about the model and potential products coming from its output.	Additional information on the NWM is available here: https://water.noaa.gov/about/nwm and products from NWS/OWP, based in part on the NWM, are available here: https://www.weather.gov/owp/operations
How does the OWP interface with NCEP?	OWP coordinates with the NCEP Weather Prediction Center (WPC) on the creation of hydro-related text and graphical forecast products, with NCEP Central Operations (NCO) on the implementation of upgrades to the NWM, and with NCEP Environmental Modeling Center (EMC) on input data sets, validation data and modeling approaches.
Soil drainage classes and water flow related to tree failure	Thank you for submitting this as a topic of interest. Additional information is needed to better understand the question.The NWM does not operate on a fine enough scale to represent soil structure/processes and stream flow as related to tree failure.
Use of AI in analysis/forecasting	While the current formulation of the NWM does not use AI or ML, such modules--like long-short term

	memory--will be included as appropriate when the NWM switches to use of the Next Generation Water Resources Modeling Framework as its underlying code base.
Is the NWM data via Microsoft available via STAC?	This dataset isn't currently available through STAC. This is a complicated dataset. Join in at https://github.com/stactools-packages/noaa-nwm if you're interested in helping.
Why does the NetCDF file format change from 2006 to 2007? All files post 2007 are encoded as NetCDF3, while 2006 and prior are encoded as hdf5. What happened to the forcing files for the 2020 model year? The folder is present, but the files inside are actually the output files.	The file formats change from 2006 to 2007 due to the application of a different processing program to package up the data before and after that date. However, the underlying version of the NWM is the same.
Interested in learning about the National Water Model output formats and contents, and how to access them easily	NWM data and related products are made available through the operational NOMADS data dissemination platform, NOAA Open Data Dissemination (NODD) cloud partners, and GIS services. The below links reflect access options and include metadata which includes details on data formats and content. For example, NOAA platforms include: https://nomads.ncep.noaa.gov/pub/data/nccf/com/nwm/prod/ https://www.weather.gov/owp/operations. https://maps.water.noaa.gov/server/rest/services/nwm Links to cloud data access points via NODD and NODD's cloud partners include: ● Azure Planetary Computer and Microsoft AI for Earth Datasets ● Google Cloud ● AWS Registry of Open Data
Options for data analysis in the cloud	There are numerous options for NOAA data analysis in the cloud. This Office Hours discussed one option via Microsoft Azure. Others include, but are not limited to:

	Amazon Web Services (AWS): AWS S3: Store data in AWS Simple Storage Service (S3) and then use AWS analytics services like Amazon Athena, Amazon Redshift, or AWS Glue for data processing and analysis. AWS Lambda: One can create serverless functions using AWS Lambda to automate data processing tasks on NOAA data. Google Cloud Platform (GCP): Google Cloud Storage: Store NOAA data in Google Cloud Storage and use BigQuery for data analysis. Google Dataprep: Use Dataprep to clean and prepare NOAA data for analysis. Google Data Studio: Create data visualizations and reports using Google Data Studio. Microsoft Azure: Azure Blob Storage: Store NOAA data in Azure Blob Storage and leverage Azure Data Lake Analytics, Azure Databricks, or HDInsight for data analysis. Azure Functions: Create serverless functions to process NOAA data as it's ingested into Azure. Jupyter Notebooks: One can set up Jupyter Notebooks in the cloud (e.g., using platforms like JupyterHub) to perform data analysis, visualizations, and model development on NOAA data. Data Analysis Libraries and Tools: One can utilize Python libraries like NumPy, Pandas, Matplotlib, and Seaborn for data analysis and visualization. Machine learning libraries like Scikit-Learn and TensorFlow can also be used for more advanced analysis. Geospatial Tools: For geospatial analysis, tools like GDAL, Fiona, and GeoPandas can be used to work with NOAA's geospatial data. There are various ESRI and Open

	Source GIS tools.
Historic streamflow data extraction for the Texas coastal region	National H D v2.0 stream reach, and then go into NWM for stream reaches which needs a GIS application. You can access and subset short and medium range channel forecast data for Version 2.1 via the above API. Please submit feedback if anyone tries it: data.ioos@noaa.gov
Data Access, Will NWM v3.0 be run as a reanalysis product?	Yes, a Feb 1979 to Jan 2023 retrospective run will be created using NWM v3.0 (mirroring that which was done for earlier versions of the model). V3.0 will also feature retro runs for oCONUS areas including the Puerto Rico, Alaska and Hawaii NWM domains.
Data management from acquisition, storage to dissemination	Thank you for submitting this as a topic of interest. Additional information is needed to better understand the question.
How does this product relate to the existing TWL forecasts that NOAA NCEP already produces in partnership with USGS? Are there plans to add waves to the tide and surge water levels?	There are a number of inundation forecasts out there from several sources, and unsure which specific ones this question refers to, but can be said that the NWM TWL forecasts are produced independently from other TWL forecasts. They are nationwide, seamless and span several forecast durations. Nationwide NWM TWL forecasts complement existing regional TWL forecasts which sometimes are higher resolution and/or more locally-specific. Waves may be added to the NWM TWL simulation in the future, if compute resources allow (cost-benefit assessment).
What irrigation applications one could make with NOAA's National Water Model?	The precipitation, evaporation, runoff, streamflow, snowpack and other fields could be useful. (note that NWM output is being used within the NWS to develop a public runoff-risk assessment product for applying fertilizers/pesticides) But it would be up to the end user to determine specific applications.
NWM v3.0 retrospective data access	NWM data and related products are made available through the operational NOMADS data dissemination platform, NOAA Open Data Dissemination (NODD) cloud partners, and GIS services. For example, NOAA platforms include: https://nomads.ncep.noaa.gov/pub/data/nccf/com/nwm/p

	[rod/](rod/) and https://www.weather.gov/owp/operations. Links to cloud data access points via NODD and NODD's cloud partners include: ● [Azure Planetary Computer](#) and [Microsoft AI for Earth Datasets](#) ● [Google Cloud](#) ● [AWS Registry of Open Data](#) Examples of how to work with the data on the cloud can be found within the above links.

QUESTION FROM DISCUSSION	RESPONSE
Is the output of the National Water Model equivalent to unimpaired hydrology? In other words, it does not account for the reservoirs, locks, and dams and other regulations, right?	When the NWM runs in operations, it includes a set of output generated by a configuration that includes reservoirs and assimilation of both streamflow and reservoir outflow obs, and also a "no-DA" or "open loop" configuration that does not include data assimilation (but still includes basic lake/reservoir objects without the benefit of DA).
Is NWM-SCHISM being used for Hawaii for compound flooding and FIM as well?	Yes, the new NWM v3.0 total water level capability currently covers Hawaii as well. FIM coverage for Hawaii is not yet available but will roll out with other domains over time (Hawaii FIM coverage currently scheduled for 2025).
Who do we contact for technical problems while retrieving NWM data? Name and email please.	For questions about retrieving real-time operational NWM data from the NOAA NOMADS server, contact ncep.pmb.dataflow@noaa.gov. For questions regarding cloud access to NWM data (real-time and retrospective data), please email nodd@noaa.gov. For general questions on the NWM, please contact Brian.Cosgrove@noaa.gov
Is the plan to include all of Alaska for the next version (in 2025)?	Alaska coverage will expand in future versions (but the next upgrade, TWL-focused v3.1 in 2025, will not feature this). That will occur in subsequent versions with specific timing TBD.

What data assimilation method are you using for streamflow and reservoir outflow?	Streamflow nudging (direct insertion) for now, with ongoing projects to improve DA.
In Sudhir's slide, does "Reanalysis" mean "Retrospective"?	Yes, on this slide 'Reanalysis' is used interchangeably with the term 'Retrospective data' (differentiating this NWM data from real-time forecasts and reforecasts).
How far back is the reanalysis data available?	NWM v2.1 retrospective data extends from 1979-2020. The forthcoming NWM v3.0 retro data set will extend from 1979-2023.
Are there plans to incorporate waves (in particular wave setup and runup) into the "TWL" and FIM?	Inclusion of wave processes (in our configuration of SCHISM) will be examined for the next version (NWM 3.1)--balancing compute with benefits to simulation results.
Where to access long term simulations (44 years retrospective?) NWM files?	NWM retro data can be accessed on cloud providers, with links on the NODD page. Realtime ops data is available on the NCO NOMADS page (and some cloud providers make realtime ops data available as well). Links to cloud data access points via NODD and NODD's cloud partners include: ● Azure Planetary Computer and Microsoft AI for Earth Datasets ● Google Cloud ● AWS Registry of Open Data NWM retrospective data AWS catalog of NWS retrospective data: https://registry.opendata.aws/nwm-archive/ V2.1 NetCDF: https://noaa-nwm-retrospective-2-1-pds.s3.amazonaws.com/index.html It is planned that the new NWM v3.0 retrospective data set will be hosted by these providers later this year.
(for clarification) So NWM v3.0 would be available in 2025 and NWM v4.0 in 2026. And the	The current operational version of the NWM is v3.0, implemented last month. We are in the process of arranging transfer of v3.0 retrospective data to the cloud

currently available latest version is NWM v2.0 through AWS (?) or Azure (?)	providers, which should occur over the next several weeks. Currently, the latest retro version they have is v2.1. The next version of the NWM is v3.1 due out in 2025 and focusing on TWL improvements. V4.0 (first version to use NextGen framework) will come out after that potentially in 2026/2027.
Is the NWS planning to use the NWM for the River forecasting in the future? If so, when can we expect that to happen?	The NWM is used in a complementary fashion within the NWS hydrology program--used in conjunction with existing sources of streamflow forecast data from the RFCs and other offices
Does the NWM v3.0 establish a two-way coupling with the SCHISM model to estimate the total water level?	Improvements in skill for v3.0 were obtained via improved calibration and regionalization of parameters, improved forcing data, and improvements to model physics (improved treatment of urban-area runoff and a new infiltration scheme).
The reduction observed in the NWM v3.0 estimates for PB is remarkable. Does NWM v3.0 incorporate modifications in the model's schemes, particularly in the routine scheme?	Currently the TWL configuration with SCHISM uses a 1-way coupling with the NWM, but this will change to 2-way once hydraulic routing is implemented in the inland NWM in a future version.
Does Microsoft replicate the NODD bucket from an Azure US zone to EUWest for Planetary Computer use?	The majority of our datasets, including goes, are in the West Europe region. Our documentation, like https://planetarycomputer.microsoft.com/dataset/storage/noaa-nwm, has the Azure region. The National Water Model data are in the East US region. For datasets cataloged with STAC, the STAC collection includes an `msft:region` property that catalogs the Azure region.
Does NWM have a root zone component and groundwater system included in the framework?	The NWM currently uses the Noah-MP LSM for its soil component, and that LSM includes a root zone. The NWM represents groundwater storage via a simple conceptual groundwater storage model (an area for future enhancements)

Has it been decided which models will be included in the main NextGen framework?	NextGen currently includes components like Noah-OWP-Modular, Conceptual Functional Equivalent (CFE), TopModel, Sacramento and snow/glacier and other models. With more being added. The full set of formulations that will be used in NextGen (and in NWM v4.0 which will be based on NextGen) have not yet been selected.
@Tom, Can you share the notebook?	Re notebooks, from CIROH developers conference: https://github.com/TomAugspurger/noaa-nwm/tree/main/content (@Tom may have an updated version...) The example notebook is at https://nbviewer.org/github/microsoft/AIforEarthDataSets/blob/main/data/noaa-nwm-example.ipynb
@ Tom, mind directing me to more info on the STACKerchunkBackend? Or the notebook you shared? Thanks!	Keep an eye on https://github.com/fsspec/kerchunk/pull/372 and https://github.com/stac-utils/xstac for where that backend will eventually end up. In the meantime, you can check the kerchunk docs for how to go from Kerchunk indices -> reference filesystem -> xarray.
Which 10% of the country is currently covered by FIM products? I see the experimental FIM products in layer list in the viewer (https://viewer.weather.noaa.gov/water), but I don't see any layers showing up on the map when I toggle them on and off	It currently covers portions of PA and NY along with Texas near Houston. Public FIMs are limited to the inland areas, not right along the coast. Data from the new coastal TWL output of the NWM will be added into FIM processing in the next 1-2 years, expanding FIM coverage to the coastline.
Does NODD plan to release kerchunk headers with NWM data in the future?	That's something we're looking at. We're exploring how we want the kerchunk index to be available but haven't decided yet how to disseminate. Doing work internally. We created a set of kerchunk headers for the 2.1 retrospective on AWS.
Are there any Github links or documentation to all these hydrologic models (e.g.	@KJ, if you would, reach out to me after the webinar over email (Brian.Cosgrove@noaa.gov) and I'll look to

Sacramento, TopModel etc.) within NextGen? I could only find CFE model within NextGen framework	see if I can provide further details. (contact has been made over email on 10/19)
Is the Total Water Level (TWL) variable stored in the NWM v3.0 netcdf output?	Yes, the NWM water level elevation output is stored in NetCDF output files (this is separate from the post-process-produced FIM maps). NWM coastal output data can be found in 'coastal' sub-directories, for example: https://nomads.ncep.noaa.gov/pub/data/nccf/com/nwm/v3.0/nwm.20231019/analysis_assim_coastal_atlgulf/
I would like to see a daily or monthly averaged dataset from the long term retrospective dataset? what is your thought?	While we don't post data that has been processed like that, I think that could potentially be a good application of cloud processing perhaps
It was hard to figure out the latest version of the NWM when I initially used it. It's confusing to find the difference between reanalysis versions between different CSPs. Can I have some clarification?	Thank you for the feedback, we will record it. We're looking at how to restructure these data. The data you're seeing is on 3 CSPs, and is identical across providers. You can determine the latest data by looking at the version number (number increases with latest data, version 2.1 newer than version 1.2). Data from NWM v3.0 will be posted to cloud providers in the next 1-2 months.
I am just starting to know about the NWM. I am needing historic streamflow data specifically to the Texas coastal region. What is your suggestion; where do I start and how do I extract the streamflow data for that specific region?	National HD v2.0 stream reach, and then go into NWM for stream reaches which needs a GIS application. You can access and subset short and medium range channel forecast data for Version 2.1 via the above API. Please submit feedback if anyone tries it: data.ioos@noaa.gov
How good are the evapotranspiration datasets? Is it good to use in agricultural water management?	We haven't conducted a verification study on flux data. This is a relatively new model (6 years old), and we're concentrating heavily on streamflow currently, so no accuracy metrics for evaporation products are available yet. The model represents different vegetation types, but not specific crops types. The land surface component of

Lots of potential for use in water management (ag). The evaporative flux datasets, when trying to look into ag system, can we use those data and how reliable is the simulation data for estimation of how crops are using the water?	the NWM is a 1km gridded scale model, and uses NOAA-MP land surface model at the moment, so looking at literature for general ag references related to NOAH-MP may be helpful (although parameters, forcing etc are likely different).
Updating calibration method in the new version– what calibration method was used in the previous version, and what will be the new calibration method? Will you use the same for nexgen outputs?	Nexgen will use some of the same approaches used in the current water model, but because it's such a different framework, it can't use the same exact approach. Right now (and probably with NextGen), we look for basins we can specifically calibrate with gauge observations, look at streamflow obs and use a dynamic (DDS) approach to minimize the error function we've chosen (KGE). Parameters from these directly-calibrated basins are then 'spread' to other basins that we are not able to directly calibrate, using a basin-similarity regionalization process.
18 hr, 10 day, and 30 day forecasts were mentioned, but what is the time increment (e.g. hourly water levels?) for each of those forecasts?	For streamflow, 18-hour and 10-day is hourly output. 30-day NWM Long-Range is temporally coarser. You can gather details of file output timing here (actual files and timing): https://nomads.ncep.noaa.gov/pub/data/nccf/com/nwm/prod/

3. Office Hours Organizing Team

Name	Title
Cindy Elsenheimer	NWS Partner Engagement Lead
Sudhir Shrestha	NWS National Water Center Technical Director Web and Data Services
Brian Cosgrove	NWS National Water Center Technical Director for National Water Model
Adrienne Simonson	NODD Director
Patrick Keown	NODD Program Manager
Jenny Dissen	NODD Engagement Lead / NCICS / NC State University
Katelyn Szura	NODD Communications Lead
Jonathan Brannock	NODD Lead Cloud Software Engineer /NC State University
Otis Brown	Director, NC Institute for Climate Studies (NCICS) / NC State University
Tom Augspurger	Geospatial Infrastructure Engineer

4. Poll Results

Poll 1		
Question	**Answer**	**Count**
How do you access NWM data today?	On-prem via NOAA	12
	Cloud	11
	Both/Either	6
	3rd Party/Web-based Viewer	4
	None/Other	24

Poll 2		
Question	**Answer**	**Count**
My primary goal for attending today is:	Technical use and access of NWM data	31
	To learn about cloud access to date (e.g. NODD Program)	15
	Meet and engage with NOAA staff scientists	8
	Learn about Microsoft Azure access and tools	6

5. Resources / References
   - NOAA Open Data Dissemination | NODD Email
   - NWS Office of Organizational Excellence | Cindy.Elsenheimer@NOAA.GOV
   - The National Water Model | National Water Center Products and Services
   - National Water Model Access via Microsoft Azure | Microsoft Planetary Computer

Thank you to our participants for engaging in this discussion!