Received: 20 January 2023 | Accepted: 11 November 2023
DOI: 10.1111/1752-1688.13184

AMERICAN WATER RESOURCES ASSOCIATION
JAWRA
Journal of the American Water Resources Association

RESEARCH ARTICLE

NOAA's National Water Model: Advancing operational hydrology through continental-scale modeling

Brian Cosgrove[1] | **David Gochis**[2] | **Trey Flowers**[3] | **Aubrey Dugger**[4] | **Fred Ogden**[3] | **Tom Graziano**[1] | **Ed Clark**[3] | **Ryan Cabell**[4] | **Nick Casiday**[3] | **Zhengtao Cui**[5] | **Kelley Eicher**[6] | **Greg Fall**[7] | **Xia Feng**[8] | **Katelyn Fitzgerald**[4] | **Nels Frazier**[9] | **Camaron George**[8] | **Rich Gibbs**[8] | **Liliana Hernandez**[9] | **Donald Johnson**[3] | **Ryan Jones**[3] | **Logan Karsten**[2] | **Henok Kefelegn**[9] | **David Kitzmiller**[1] | **Haksu Lee**[5] | **Yuqiong Liu**[5] | **Hassan Mashriqui**[1] | **David Mattern**[9] | **Alyssa McCluskey**[4] | **James L. McCreight**[10] | **Rachel McDaniel**[3] | **Alemayehu Midekisa**[8] | **Andy Newman**[4] | **Linlin Pan**[4] | **Cham Pham**[1] | **Arezoo RafieeiNasab**[4] | **Roy Rasmussen**[4] | **Laura Read**[4] | **Mehdi Rezaeianzadeh**[9] | **Fernando Salas**[3] | **Dina Sang**[8] | **Kevin Sampson**[4] | **Tim Schneider**[4] | **Qi Shi**[9] | **Gautam Sood**[1] | **Andy Wood**[4] | **Wanru Wu**[1] | **David Yates**[4] | **Wei Yu**[4] | **Yongxin Zhang**[4]

[1]Office of Water Prediction, NOAA National Weather Service, Silver Spring, Maryland, USA
[2]Airborne Snow Observatories, Inc., Mammoth Lakes, California, USA
[3]National Water Center, NOAA National Weather Service, Office of Water Prediction, Tuscaloosa, Alabama, USA
[4]Research Applications Lab, NSF National Center for Atmospheric Research, Boulder, Colorado, USA
[5]Lynker, Silver Spring, Maryland, USA
[6]University of Minnesota, Minneapolis, Minnesota, USA
[7]Office of Water Prediction, NOAA National Weather Service, Chanhassen, Minnesota, USA
[8]Cooperative Institute for Research to Operations in Hydrology, Tuscaloosa, Alabama, USA
[9]Lynker, Tuscaloosa, Alabama, USA
[10]Cooperative Programs for the Advancement of Earth System Science

Abstract

The National Weather Service (NWS) Office of Water Prediction (OWP), in conjunction with the National Center for Atmospheric Research and the NWS National Centers for Environmental Prediction (NCEP) implemented version 2.1 of the National Water Model (NWM) into operations in April of 2021. As with the initial version implemented in 2016, NWM v2.1 is an hourly cycling analysis and forecast system that provides streamflow guidance for millions of river reaches and other hydrologic information on high-resolution grids. The NWM provides complementary hydrologic guidance at current NWS river forecast locations and significantly expands guidance coverage and water budget information in underserved locations. It produces a full range of hydrologic fields, which can be leveraged by a broad cross section of stakeholders ranging from the emergency responder and water resource communities, to transportation, energy, recreation and agriculture interests, to other water-oriented applications in the government, academic and private sectors. Version 2.1 of the NWM represents the fifth major version upgrade and more than doubles simulation skill with respect to hourly streamflow correlation, Nash Sutcliffe Efficiency, and bias

Paper No. JAWR-23-0005-P of the *Journal of the American Water Resources Association* (JAWR).
Discussions are open until six months from publication.

(CPAESS), University Corporation for Atmospheric Research, Boulder, Colorado, USA

Correspondence
Brian Cosgrove, Office of Water Prediction, NOAA National Weather Service, Silver Spring, MD, USA.
Email: brian.cosgrove@noaa.gov

reduction, over its original inception in 2016. This paper will discuss the driving factors underpinning the creation of the NWM, provide a brief overview of the model configuration and performance, and discuss future efforts to improve NWM components and services.

KEYWORDS
hydrologic modeling, flooding, decision support systems, surface water hydrology, hydrologic cycle, rivers/streams, operational modeling, land surface modeling

Research Impact Statement

The National Water Model is a high-resolution, continental-scale, operational hydrologic model providing streamflow and other guidance for a broad range of applications.

1 | INTRODUCTION

The roots of the National Water Model (NWM) can be traced back to hydrologic forecasting activities which have been an important component of National Weather Service (NWS) operations for many decades. Supporting the core NWS mission of the "protection of life and property and enhancement of the national economy", these forecasts seek to provide information on the serious riverine floods, flash floods, coastal floods and droughts which impact the United States each year (https://www.nws.noaa.gov/mission.php).

Floods are devastating natural disasters, causing billions of dollars of damage each year and putting many lives in danger (Ashley & Ashley, 2008; Zhou et al., 2018). With the exception of excessive heat, flooding leads to more weather-related fatalities on average than any other cause (NWS, 2022c). Given this, accurate and timely predictions of floods are essential. Unfortunately, the nature of these events makes them quite difficult to monitor and predict. Flash floods feature a fast onset, less than 6h from the causative event (NWS, 2002), are local in scope, and depend greatly on fine-scale weather and land surface conditions. While slower-developing, riverine floods feature their own prediction challenges, including their potentially large spatial scale, with interrelated impacts across emergency management, water management, infrastructure and forecasting boundaries.

As dangerous as too much water in the wrong place or at the wrong time can be, a paucity of water can ultimately impact a broader area for a longer duration than floods, leading to more extensive and persistent, if less obvious, damage. Touching practically every corner of society, from water management, to power and industrial production, to agriculture and recreation, drought has inflicted billions of dollars in damage to the U.S. economy since 1980 (NOAA NCEI, 2022).

Given the severe impacts of these water extremes, NWS forecasters have developed a variety of modeling tools to produce the hydrologic analyses and forecasts that are essential to ensuring public safety and optimal resource management. One of the mainstays in operational forecasting is the Community Hydrologic Prediction System (CHPS) (Restrepo et al., 2010). However, as skillful as lumped modeling approaches like those in CHPS are, they are limited by the fact that they only provide information at the outlet of the smallest basin modeling unit and cannot provide guidance inside of each basin, reflecting the highly variable land surface and meteorological conditions that impact spatially distributed flooding (Paudel et al., 2011; Reed et al., 2004). An alternative to lumped modeling is spatially-distributed modeling, which can provide streamflow and land surface information at any grid point within the model domain (Paudel et al., 2011).

Some spatially distributed gridded models used by NWS forecasters inform flash flood prediction operations indirectly—via feeding into the Flash Flood Guidance (Sweeney, 1992) postprocessor, which calculates the amount of rain needed for streams to reach bankfull conditions. Other gridded models, like the Flooded Locations and Simulated Hydrographs (FLASH) system (Gourley et al., 2017) produce simulations of streamflow and surface runoff for direct characterization of flash flood conditions. Optimized for speed, FLASH contains a simplified set of physics geared towards warm-season riverine flash flooding. The model is used at NWS Weather Forecast Offices (WFOs) and demonstrates the power of distributed modeling in short lead time flash flooding situations.

To cover longer-lead times, anthropogenic influences and snow-melt-driven flooding, River Forecast Centers (RFCs) leverage a range of modeling tools to produce streamflow forecasts for approximately 3800 locations across the United States (NWS, 2022a). These vital forecasts are used to inform emergency response, optimize power generation and commercial navigation, and further water management activities. When distributed across the country, however, even this seemingly large number of forecast locations (3800) leaves wide swaths of the domain without forecast coverage, with only 100,000 out of 3,500,000 miles of river covered by a forecast. These spatial gaps include a wide variety of areas such as power generation sites, ecologically sensitive regions, cities and recreation areas—locations where guidance on water flow is critical. Variation in approach also handicaps development efforts, as there is no single, uniform nationwide system which can act as a unified platform for model development.

Lastly, the current suite of regional RFC hydrologic models are manually labor-intensive and do not represent the full set of coupled inland-coastal-cryosphere-groundwater processes necessary to fully capture the hydrologic cycle. While the streamflow forecasts from these models are, in and of themselves, highly useful, the holy grail for emergency responders and many end users is flood inundation—namely the extent of the ground which is covered by flood waters. Without direct knowledge of this quantity via maps of surface inundation for all portions of their area of interest, end users lack key, actionable intelligence allowing them to understand and react to an ongoing or forecast flood event.

Against this backdrop, the National Academy of Sciences (NAS) issued a report in 2012 with guidance on the improvements needed in the NWS hydrologic forecasting enterprise (National Research Council, 2012). A sampling of the critical recommendations includes:

- Implement a new hydrologic modernization effort.
- Improve pathways for collaboration and accelerate research to operations efforts.
- Transition RFC forecasters away from the execution of labor-intensive models and towards an “over the loop” role, enabling a shift in focus to model and product development, forecast interpretation, and decision support.
- Add value to hydrologic forecasts through the use of more advanced models, data assimilation (DA) and employment of more sophisticated ensemble techniques.

Combined with the previously noted forecast system shortcomings, these NAS recommendations firmly established the outlines for a new NWS operational hydrologic forecasting capability—the NWM. These outlines were translated into a physical system via a timely combination of five main resources:

1. The Integrated Water Resource Science and Services (IWRSS) consortium—consisting of Federal Water Agency partners (NOAA, USGS, USACE, FEMA) and seeking to develop shared plans for a virtual or physical center to advance water resources prediction nationwide;
2. The new National Water Center in Tuscaloosa, Alabama—acting as a catalyst and collaborative focal point for hydrologic research and operations support (Title III of Public Law No: 116-271—Coordinated Ocean Observations and Research Act of 2020, 2020);
3. The creation of the Weather Research and Forecasting (WRF) Hydrological Modeling Extension (WRF-Hydro; Gochis et al., 2020) at the National Centers for Atmospheric Research (NCAR)—providing a flexible, community-based hydrologic modeling foundation;
4. The NWS-CUAHSI Summer Institute program at the National Water Center—providing a testbed for continental-scale real-time hydrologic modeling; and
5. Congressional programmatic mandates to improve hydrologic forecasting—providing the resources needed to accomplish hydrologic modeling and product modernization

These driving forces provided the comprehensive mix of science, programmatic, staffing and funding resources needed to create the operational NWM hydrologic modeling system. This system delivers guidance which complements the essential and skillful forecasts already produced by RFCs, as well as addresses the needs of an expanded set of end users and critical mission areas.

2 | OPERATIONAL MODEL CONFIGURATION

The operational configuration of the NWM was driven by the wide variety of end user needs including flood and drought forecasting, water resource management and emergency response. Several configurations of the NWM are run operationally on the NOAA Weather and Climate Operational Supercomputing System (WCOSS) computing platform over the range of modeling domains depicted in Figure 1. These configurations include: (1) A simulation of the past several hours using the best available observed forcings and streamflow DA to establish initial conditions for forecasting, which we call the “CONUS Analysis and Assimilation” cycle, (2) CONUS Short-Range 18 h deterministic forecast, (3) CONUS Medium-Range 10 day ensemble forecast, (4) CONUS Long-Range 30 day ensemble forecast, (5) Hawaii and Puerto Rico/USVI Analysis and Assimilation current snapshot, and (6) Hawaii and Puerto Rico/USVI 48 h short range forecast (Table 1). Open-loop (no DA) configurations of the Analyses, CONUS medium-range, and island short-range forecasts are executed as well.

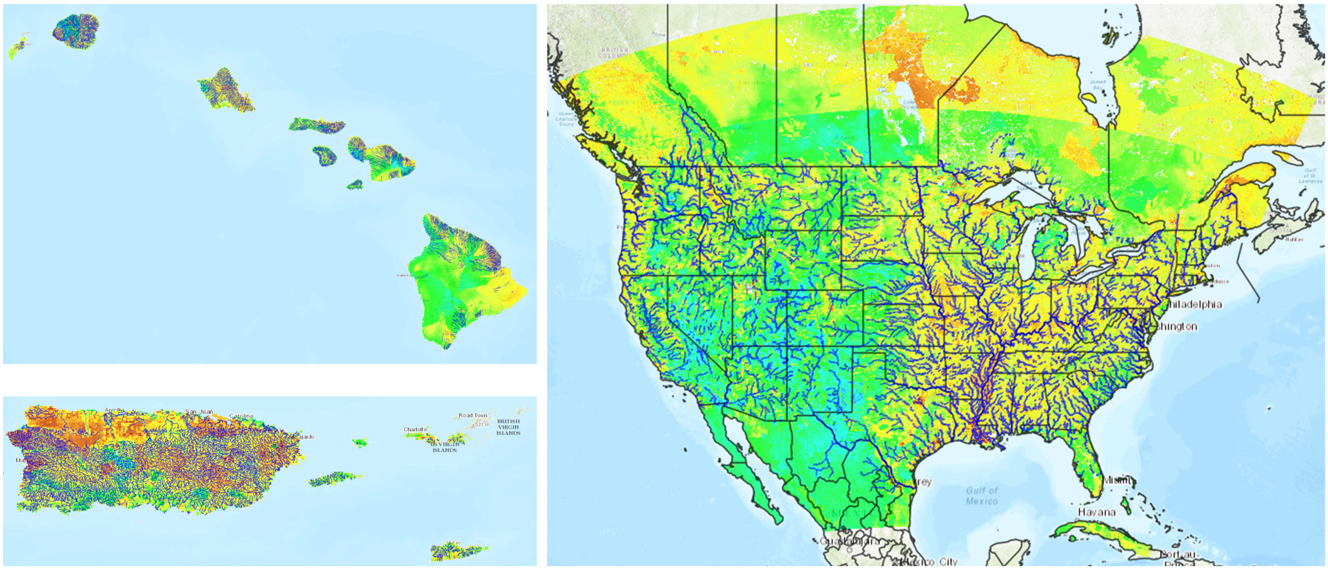


FIGURE 1 National Water Model (NWM) v2.1 operational domains including CONUS (right), the Hawaiian Islands (upper left) and Puerto Rico and the US Virgin Islands (lower left). Blue lines are a thinned depiction of NWM streamflow channels, while shading indicates land surface modeling domain.

TABLE 1 Description of each NWM v2.1 modeling configuration, with abbreviations as analysis and assimilation (AnA), Short-Range Forecast (SRF), Medium-Range (MR), and Long-Range Forecast (LR).

NWM v2.1 model configuration	Simulation length	Cycling frequency	Atmospheric inputs	Additional detail
CONUS Extended AnA	28 h lookback	1× day	HRRR, RAP, MRMS, MPE	Features data assimilation
CONUS AnA	3 h lookback	24× day	HRRR, RAP, MRMS, MPE	Features data assimilation
CONUS SR	18 h forecast	24× day	HRRR, RAP	
CONUS MR	240 h forecast	4× day	GFS	7 ensemble members
CONUS LR	30 day forecast	4× day	CFS	4 ensemble members
Hawaii AnA	3 h lookback	24× day	NAM-Nest, MRMS	Features data assimilation
Hawaii SR	48 h forecast	2× day	WRF-ARW, NAM-Nest	
PR/USVI AnA	3 h lookback	24× day	NAM-Nest, WRF-ARW	Features data assimilation
PR/USVI SR	48 h forecast	2× day	WRF-ARW, NAM-Nest	

Abbreviations: CFS, Climate Forecast System; GFS, Global Forecast System; HIRESW WRF-ARW, High-Resolution Window Forecast System Advanced Research Weather Research and Forecasting; HRRR High-Resolution Rapid Refresh; MPE, Multisensor Precipitation Estimator; MRMS, Multi-Radar Multi-Sensor; NAM-Nest, North American Model High-Resolution Nest; RAP, Rapid Refresh; WRF, Weather Research and Forecasting.

2.1 | Continental United States configurations

The accuracy of hydrologic forecasts are heavily influenced by the quality of the model initialization (Li et al., 2009). As such, the various analysis and assimilation (AnA) configurations which initialize each NWM forecast play a critical role in the modeling process. Over the United States, these initial conditions come from a set of three separate AnA cycles—two linked to the NWM short- and medium-range forecasts, and one linked to the NWM long-range forecast. Each of these analysis cycles is started from the previous AnA iteration, making them continuously linked rather than event-based in nature, and eliminating the need for spin-up simulations. Further details of these configurations are given in Appendix A.

The Short-Range Forecast configuration produces guidance for forecasters and emergency responders who deal with events such as flash floods and other situations which can rapidly evolve based on changing meteorological and hydrologic conditions. It cycles hourly, is forced with meteorological data from the High-Resolution Rapid Refresh (HRRR) and Rapid Refresh (RAP), is initialized from the latest AnA solution, and produces hourly deterministic forecasts of streamflow and hydrologic states out to 18 h.

The Medium-Range Forecast configuration which is executed four times per day at 00Z, 06Z, 12Z, and 18Z, is initialized from the latest AnA solution and is forced with Global Forecast System (GFS) model output. Unlike the deterministic Short-Range configuration, the

Medium-Range configuration produces ensemble guidance. Member 1 of this configuration is forced with model output from the most recent GFS NWP model run, and extends out to 10 days. While starting from the same set of initial conditions as member 1, members 2–7 ingest time-lagged forcing data from successively older forecast cycles of the GFS. These six members extend to 8.5 days, creating an overall seven member equal-weighted medium range ensemble forecast with an 8.5 day forecast horizon. At 30-days in length, the longest set of NWM forecasts is produced by the Long-Range Forecast. As with the Medium-Range configuration, the Long-Range configuration is a forcing-based ensemble which cycles at 00Z, 06Z, 12Z, and 18Z. However, it differs in that it consists of four members which are each forced with a different Climate Forecast System (CFS) forecast member from a single CFS model cycle.

Adding to the standard configurations outlined above, several open-loop or "no-DA" configurations are executed as well. These open loop analyses lack the assimilation of USGS stream gauge observations that occur in the standard analysis, and provide end users with insight into the native model behavior, unaffected by observation-based corrections. These open loop analyses are also used to initialize corresponding CONUS medium-range, and island short-range forecasts.

All of the analyses and forecasts are executed with the same 1 km resolution land surface grid, and use the same NHDPlus-based hydrofabric supporting routed channel output at over 2.7 million stream reaches. Additionally, the Analysis, Short-, and Medium-Range simulations rely on identical 250 m CONUS routing grids to perform surface and sub-surface (non-channelized) routing.

2.2 | Hawaii and Puerto Rico/USVI configurations

The NWM v2.1 operational Hawaii and Puerto Rico/USVI configurations feature several important differences from the CONUS configurations discussed above. These differences are driven by the selection of observation- and model-based forcing data which is available for these island domains. The most significant difference is that only Analyses and Short-Range forecasts are produced—the latter extending out to a 48-h forecast horizon. Further, whereas the CONUS Analysis and Short-Range cycles are driven in part by data from the HRRR and RAP models, the island configurations draw meteorological input from (1) Shortwave and Longwave Radiation: North American Model Nest (NAM-Nest; Carley et al., 2017; Rogers et al., 2017) and (2) Remaining Fields: WRF-ARW models. The final difference lies in the spatial resolution of the Hawaii and Puerto Rico NWM configurations. While they rely on the NHDPlus-based hydrofabric and feature the same 1 km base land surface grid resolution, the surface and subsurface routing grid is set at 100 m, versus the 250 m resolution used over the CONUS, reflecting the high topographic relief of the islands. As occurs in operations over the CONUS, real-time USGS streamflow observations are assimilated into the Island Analysis configuration. Additionally, MRMS precipitation is ingested as forcing into the Hawaii Analysis, while the Puerto Rico/USVI Analysis relies upon the HIRESW—with both island Analysis configurations using precipitation from the NAM-NEST when their primary sources are unavailable. The remaining Analysis forcing fields are obtained from the NAM-NEST.

3 | NWM STRUCTURAL OVERVIEW

Functioning across the variety of domains and forecast configurations described above, the NWM is a multi-faceted modeling system as opposed to a singular hydrologic process model. It consists of the community WRF-Hydro modeling system, a detailed geospatial-hydrofabric set of data layers, a Meteorological Forcing Engine (MFE) and an automated model calibration system. Each of these components is flexible and extensible, with structural support for the addition of new modular process modeling capabilities.

WRF-Hydro is a community modeling framework that was initiated in 2003 as the 'Noah-distributed' modeling system (Gochis & Chen, 2003) and encompasses several modules and datasets, which together allow for efficient and robust, continental-scale, high-resolution simulations of the Nation's hydrologic systems. Representations of infiltration, soil hydraulic physics, evapotranspiration, snowpack accumulation and ablation, lateral overland and subsurface flow, baseflow and river channel flow exist inside the model within various model components. These processes can each run on different spatial frameworks (e.g. rectilinear grids, catchments, river reach vectors and reservoir objects) thus providing flexibility in process representation and computational efficiency. The multi-scale WRF-Hydro system has been developed for application on large clusters and high performance computing systems so is applicable to large-domain, high-resolution simulations. It is currently used for a wide range of hydrometeorological research and operational forecasting applications (Arnault et al., 2021; Givati et al., 2016; Lahmers et al., 2019; Pal et al., 2020; Rummler et al., 2019; Senatore et al., 2015, 2020; Verri et al., 2017; Xiang et al., 2017; Yucel et al., 2015).

3.1 | NWM hydro-geo fabric

Each of the components within WRF-Hydro are supported by a detailed hydro-geospatial fabric of inputs—the means by which the modeling system represents the geometry and topology of river networks and the description of the land surface and shallow subsurface at scales appropriate for NWM simulations and end-use applications. The NWM V2.1 modeling domain extends from central Mexico to central Canada, and includes the major Hawaiian Islands, Puerto Rico, and the U.S. Virgin Islands (Figure 1). Land surface processes are modeled using a 1 km gridded resolution over all domains, with surface overland flow and saturated sub-surface flow routed on a 250 m grid over the CONUS, and on a 100 m grid over Hawaii, Puerto Rico, and the U.S. Virgin Islands. In the NWM configuration of WRF-Hydro, channel routing of run-off occurs on a vector-type network of channel reaches (Table 1). The NWM channel routing network is largely derived from the National Hydrography Dataset Plus Version 2 (NHDPlus; McKay et al., 2012), which is based on an integration of the medium-resolution National Hydrography Dataset (NHD), National Elevation Dataset (NED), and National Watershed Boundary Dataset (WBD). While a high-resolution version of NHDPlus has become available, the NWM leverages the 1:100,000 scale flowlines, catchments, waterbodies, gauge associations, and value-added attributes of the Medium Resolution NHDPlus for operations, providing streamflow output for over 2.7 million river reaches nationwide.

While the NHDPlus Version 2.1 dataset underpins the vast majority of the features within the NWM hydrofabric, a number of modifications were necessary to support the spatial requirements of the NWM's extensive domain. The large catchments representing upstream areas in Canada and Mexico were replaced with hydrography and elevation-derived catchments roughly equivalent in scale to NHDPlus. Additionally, divergent flow paths were eliminated, Strahler stream orders re-calculated, and a continuous flow network was ensured by examining all interior network endpoints and reconnecting the network wherever possible. In this way, a seamless computational river network was constructed which could support operational hydrologic needs across the country. In areas tributary to the U.S., the channel networks and catchment delineations were custom-derived and linked to the NHDPlus networks. In areas beyond tributary regions, no channel routing is performed in the NWM. The USGS-World Wildlife Federation 90 m 'Hydrosheds' topography dataset was used in these regions where NHDPlus data was unavailable (Lehner et al., 2008).

3.2 | Physics processes

3.2.1 | Column land surface physics

The core physics of the multi-scale WRF-Hydro modeling system consists of multiple process modules including a column land surface model, modules for overland and saturated subsurface lateral flows, a bucket-type parameterization for baseflow, a channel flow module, and a reservoir routing module. The column land surface model used in the NWM is an updated version of the community NoahMP land surface model (Niu et al., 2011; Yang et al., 2011). NoahMP is itself a multi-physics model with numerous process representation options for the accounting of canopy energy exchange, snowpack, runoff and infiltration, groundwater, dynamic vegetation, surface aerodynamic exchange and land cover description. The time-step for the execution of NoahMP is identical to that of the meteorological forcing input which is 60 min.

3.2.2 | Overland flow formulation

While the NoahMP land surface model is configured to run on a 1 km rectilinear grid, the overland and saturated subsurface flow routines execute on a 250 m CONUS (100 m over the Hawaii and Puerto Rico/Virgin Island domains) rectilinear grid that is forced to edge-match the 1 km grid. Infiltration-capacity excess from NoahMP which exceeds a pre-defined retention depth value is allowed to flow overland using an option within WRF-Hydro that is an explicit formulation of the diffusive wave simplification of the full St. Venant equation for overland flow, similar to Julien et al. (1995; Gochis et al., 2020). The one-dimensional steepest descent gradient estimation option within WRF-Hydro is used in the NWM to minimize forecast latency. To satisfy Courant constraints in the diffusive wave formulation, when executed at a 250 m grid spacing an overland flow routing time step of 10 s is used. Overland flow, with depths exceeding the specified retention depth, and intersecting 250 m grid cells that have been designated as 'channel' grid cells, is sequestered as channel inflow. Overland flow-derived channel inflow for each NHDPlus channel reach is aggregated from all channel grid cells within each NHDPlus catchment associated with each NHDPlus channel reach. We note here that the 250 m overland and subsurface (described below) routing functions are used in all NWM configurations except the long-range, 30-day forecast configuration due to computational constraints in completing the long-range forecasts. Instead of performing the high-resolution overland and subsurface routing, the long-range configuration simply aggregates surface runoff and base-flow across NHDPlus catchments and the stream inflow contributions. This more simplistic formulation greatly reduces computational expense in forecast production but both reduces the kind of outputs produced by the model and also requires a separate calibration due to differing process representations.]

3.2.3 | Saturated subsurface flow formulation

Similar to overland flow, shallow saturated subsurface flow is executed on a 250 m rectilinear grid (100 m over the island domains). A Boussinesq approximation is employed as in Wigmosta et al. (1994) which results in an effective 2-dimensional calculation of saturated subsurface lateral transport. In versions 1.0–2.1 of the NWM, subsurface soil saturation characterization is determined from the 2 m deep, 4-discretization soil moisture representation in NoahMP. In this soil column, the water table is defined as the top-most discretization which features saturated soil. As such, the subsurface flow routine in WRF-Hydro does not represent perched lateral transport. The 60-min time-step of saturated lateral flow is synchronized with the execution of the NoahMP time step. Soil moisture accounting following lateral transport can result in full column saturation and exfiltration of soil water to the land surface where it may subsequently participate in overland flow. Numerical details of the saturated subsurface flow formulation are provided in Gochis et al. (2020).

3.2.4 | Baseflow formulation

Deeper groundwater contributions to streamflow are represented using a relatively simplistic, non-linear, conceptual baseflow bucket model. The spatial unit of the baseflow buckets is that of the uniquely-defined NHDPlus catchments. Inflow into the buckets is derived from drainage through the NoahMP soil column and is calculated as a scaled free-drainage boundary flux. A spatial weighting of NoahMP grid cells encompassed within the NHDPlus catchments is used to estimate which grid cells and fractions of grid cells contribute soil column drainage to each bucket. The baseflow bucket then discharges baseflow to the upstream node of the corresponding NHDPlus stream segment associated with each NHDPlus catchment using an empirical exponential storage-discharge formulation, where the discharge from the bucket to the river channel is specified as a function of the fractional storage capacity of the bucket.

3.2.5 | Channel flow

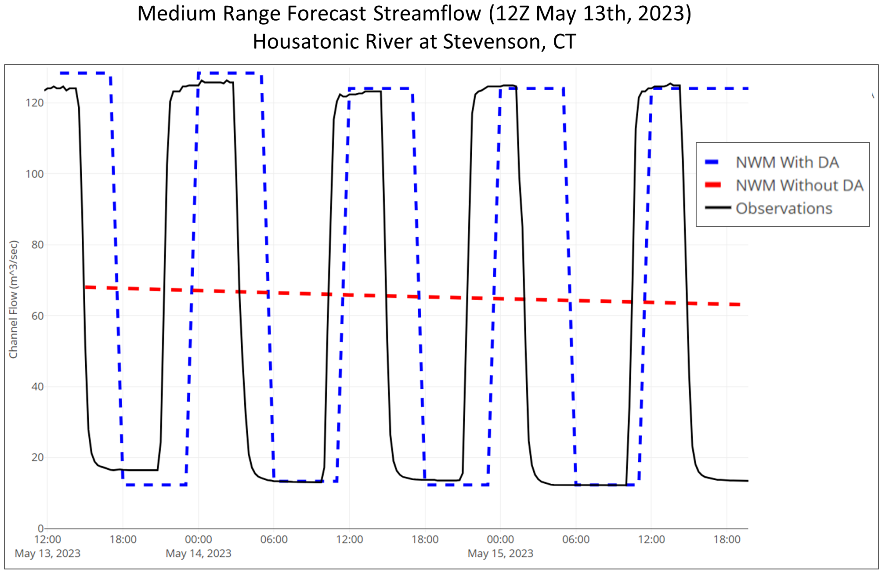
Channel inflow components coming from overland and baseflow are ingested into the channel at the upstream node of the relevant NHDPlus channel reach. Channel flow is then estimated using the iterative Muskingum-Cunge hydrologic routing formulation adapted from Chow (1959) as described in detail by Read et al. (2023), and flows are routed down the NHDPlus-defined channel and reservoir network.

3.2.6 | Reservoir accounting

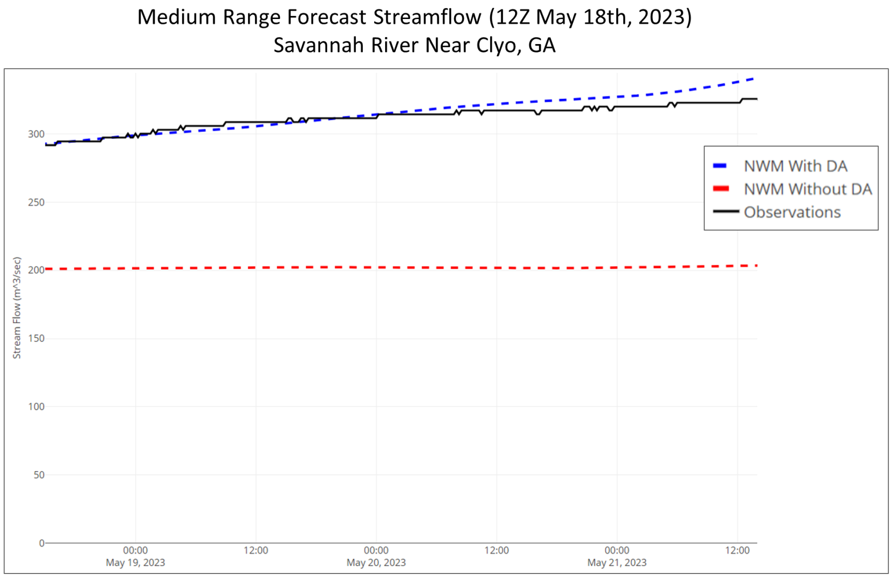
Given the thousands of reservoirs spread across the country, and the degree to which they can impact streamflow across all flow regimes, accurate representation of reservoirs is critical for NWM analyses and forecasts. More than 5000 of these reservoirs are already represented in the NWM, but in a fairly simplistic fill-and-spill manner. This basic approach does not mimic the real-world manual operation of reservoirs, and instead uses fixed orifice discharge and reservoir spillway estimation methods. Given the impractical nature of obtaining and applying thousands of operating functions, a new module was developed in NWM v2.0 to simulate the anthropogenic impacts of reservoirs with a two-pronged automated approach for use at several hundred of these locations. The approach uses a persistence-type treatment whereby USGS and USACE-supplied streamflow measurements at, or immediately downstream of, reservoirs are used to correct NWM reservoir states and releases. Additionally, RFC-supplied observations and forecasts of reservoir outflow are assimilated into the NWM for use in the NWM analysis and short- and medium-range forecasts (Figure 2a). Future versions of the NWM will leverage this flexible ingest mechanism to assimilate RFC-sourced observations and forecasts of diversion-related flow.

3.3 | Real-time streamflow DA

The NWM has assimilated real-time streamflow observations, specifically volumetric flow rates, since its v1.0 inception. The purpose of assimilating streamflow observations is to help mitigate the impact of background model error propagating downstream of observed flow locations in the NWM's analysis cycle and on its forecast initial conditions (Figure 2b). The value of the NWM streamflow DA in reducing model forecast errors increases proceeding downstream on streams and rivers which have progressively more gauges, due to the influence the gauges have on conditioning upstream flows.



(a)



(b)

FIGURE 2 Impact of assimilating River Forecast Center (RFC)-supplied reservoir outflow data (a; left) and USGS streamflow observations (b; right) on NWM v2.1 medium-range streamflow forecasts.

A simplified, computationally-efficient nudging DA was selected for use in the operational NWM. Real-time streamflow observations from the USGS NWIS data service are quality screened and directly assimilated, or inserted, into the analysis cycle of the model. Assimilation temporal weighting factors are applied to the observations, which relax the model solution back to its open-loop, or unassimilated, value. Streamflow observations deemed to be of high quality that were recorded within 15 min of the model analysis time are given full weight, while observations between 15 min and 2 h old are given proportionately lower weights. Observations older than 2 h are not used. The NWM also relies on the data quality and error flags accompanying real-time USGS streamflow observations. Generally, for well-managed gauging sites with long periods of record and well-established rating curves streamflow observations with high quality rankings can be considered reliable estimates of streamflow. It is acknowledged that there may be instances where site and flow dynamics may be evolving disproportionately and may not be detected until future stream surveys. In such instances, there may be additional uncertainty in the model. Similarly, during the forecast cycles of the NWM, the influence of the observation in the forecast is relaxed as the model iterates away from its initial state. At this initial state, the model flow solution is forced to match the observed flow solution, with the influence of the streamflow observation decaying to zero by 2 h into the forecast. Streamflow observations are applied only on the local stream reach where the observation is made, and then allowed to propagate under the model's flow solution downstream from the observation location. Streamflow DA is underpinned with data drawn from the USGS NWIS data service. While the NWM is able to ingest data from over 8000 gauge sites, the availability of observations varies with time, with many fewer gauge discharge values available during the winter due to frozen river conditions.

3.4 | Meteorological forcing prescription

The NWM is not interactively coupled to a parent atmospheric model. Rather, it ingests meteorological forcing from a variety of external sources including Multi-Radar Multi-Sensor (MRMS, Zhang, Howard, et al., 2016) radar-gauge and Stage IV Multisensor Precipitation Estimator (MPE) observed precipitation data, HRRR (Dowell et al., 2022), RAP (Benjamin et al., 2016), NAM-Nest, High-Resolution Window Forecast System Advanced Research Weather Research and Forecasting (HIRESW WRF-ARW; Skamarock et al., 2008), GFS (NWS, 2022b) and CFS (Saha et al., 2014) Numerical Weather Prediction (NWP) forecast data. Each of these forcing sources is ingested into the NWM MFE which is a set of unique configurations of the more general WRF-Hydro MFE (https://github.com/NCAR/WrfHydroForcing). The NWM MFE performs source data ingest, variable and units standardization, spatial interpolation, downscaling, statistical bias correction, product layering, and file output. The NWM requires seven meteorological variables; precipitation rate, 2-m air temperature, 2-m specific humidity, 10 m wind speed, surface pressure, incoming shortwave radiation, and incoming longwave radiation. Because all of the meteorological variables come from gridded analyses or other models whose grid structures are different from the NWM, each variable is interpolated to the 1 km NWM land surface model grid. In addition to spatial interpolation, several variables are subjected to downscaling and/or statistical bias correction within the MFE. When downscaling is needed, temperature, humidity and pressure terms are downscaled using climatological lapse rate adjustments while incoming shortwave radiation is downscaled using terrain adjustments to incoming solar radiation for terrain slope and aspect variations (terrain shading is not performed) on the 1 km model grid (Garnier & Ohmura, 1968; Zangl, 2005). A description of each variable, its required units and the processing operations applied to it are provided in Table 2. The MFE also applies a preferential layering approach to incoming precipitation datasets, the details of which are provided in Appendix A.

The Long Range configuration of the NWM utilizes statistically-processed model output from the CFS as forcing. The availability of a long reforecast archive of the CFS system provides a large, stable climatological database that makes the enhanced bias correction of CFS data possible. NCEP CFSv2 outputs covering the full set of required NWM forcings are first post-processed to correct biases in the raw CFSv2 data, and then downscaled and disaggregated to the standard space–time resolution of NWM MFE inputs. This bias correction approach is based on the quantile mapping (Panofsky & Brier, 1958; Wood et al., 2002) of CFSv2 outputs to match the North American Land Data Assimilation (NLDASv2) climatology (Cosgrove et al., 2003; Xia et al., 2012). Parametric quantile mapping is used for all variables, with a Dirac delta function to model instances of zero precipitation, and a Weibull distribution to model non-zero precipitation. A Gaussian distribution is used for 2-m temperature, surface pressure, and downward longwave radiation; a gamma distribution for specific humidity and wind speed; and a mean scaling of non-zero shortwave radiation. The parameters of the distributions are estimated on the CFSv2 grid and varied for all grid points across forecast hours for each specific initialization time, with a centered ±2-day moving window applied to increase sample sizes. This bias-correction procedure yields four cycles per day of bias-corrected CFS data which then force the 6-hourly NWM Long-Range ensemble.

TABLE 2 Description of each NWM forcing variable, required units and processing operations applied for the AnA, Short-Range Forecast, Medium-Range Forecast (MRF), and Long-Range Forecast (LRF) configurations.

Variable	Units	Downscaling	Bias correction
2 m air temperature	K	CONUS: MRF, LRF OCONUS: HI-AnA & SR, PR-AnA & SR	CONUS: SRF, MRF, LRF
2 m specific humidity	kg/kg	CONUS: MRF, LRF OCONUS: HI-AnA & SR, PR-AnA & SR	CONUS: LRF
Surface pressure	hPa	CONUS: MRF, LRF OCONUS: HI-AnA & SR, PR-AnA & SR	CONUS: LRF
Wind speed	m/s		CONUS: MRF, LRF
Incoming shortwave radiation	W/m^2	CONUS: MRF, LRF OCONUS: HI-AnA & SR, PR-AnA & SR	CONUS: SRF, MRF, LRF
Incoming longwave radiation	W/m^2		CONUS: SRF, MRF, LRF
Precipitation rate	mm/s		CONUS: LRF

3.5 | Model parameter specification

As a process-based, high resolution, spatially-distributed hydrological modeling system, the NWM has a very large number of model parameters which must be specified. The parameter specification process has three primary steps:

1. Initial parameter specification.
2. Parameter calibration.
3. Parameter regionalization.

Initial parameter specification is performed during model setup using landscape classification information for soils, vegetation, land use, channel hydrography, lakes and reservoirs. The datasets for these landscape classifications come from a variety of data sources which are listed in Table 3. Initial model parameter estimates are spatially-mapped to the NWM using these classifications and model parameter tables which associate model parameters with a specific classification (e.g. leaf area index, soil hydraulic conductivity, soil porosity, canopy roughness heights, etc). There currently is no sub-grid land cover/vegetation or soils type classification within the NWM, and each model grid cell has a unique classification and set of parameter values.

TABLE 3 Land surface datasets used in NWM v2.1.

Parameters	Data source
Elevation	CONUS: National Elevation Database OCONUS: HydroSHEDS
Channel and waterbody data	CONUS: NHDPlusv2 Medium Resolution OCONUS: Modified Great Lakes Hydrography Dataset (GLHD) and Elevation-Derived Hydrography
Soil type	CONUS: STATSGO Hawaii: SoilGrids Puerto Rico/USVI: gSSURGO
Land cover	CONUS: NLCD 2016 OCONUS: MODIS, Ontario Provincial Land Cover
LAI and greenness fraction	MODIS-based Monthly Climatologies

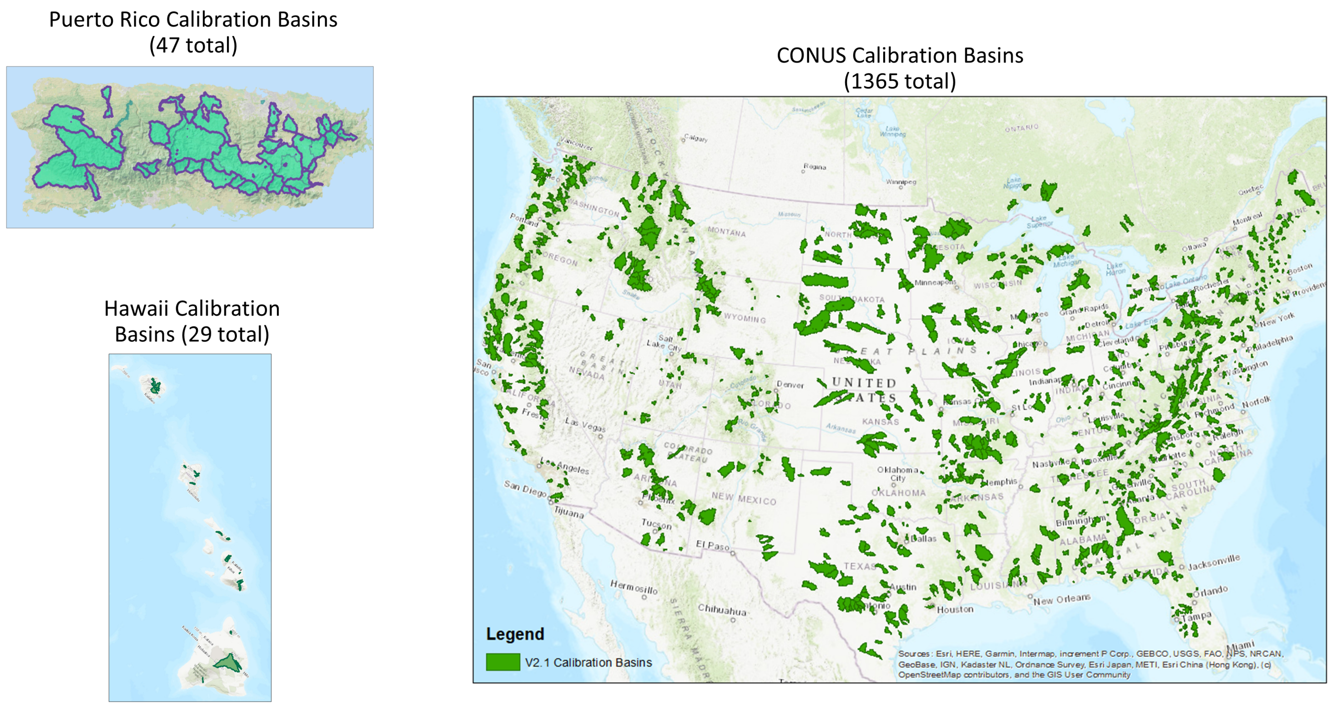


FIGURE 3 Depiction of basins directly calibrated using streamflow observations in NWM v2.1.

3.6 | Calibration and regionalization

The significant physics and parameter upgrades inherent in each version of the NWM necessitate recalibration of the NWM before each version enters into NWS operations. The calibration period for the CONUS domain of the current operational version of the NWM (v2.1) extends from October 2008 through September 2013, with a validation period spanning October 2013 through September 2016, and 2007 used for spin-up. Meteorological forcing is drawn from the Analysis of Record for Calibration (AORC) dataset developed by the NWS OWP (Fall et al., 2023). To avoid complications introduced by water management activities, direct calibration is limited to 1365 lightly regulated basins selected from USGS, California Department of Water Resources (CADWR) and Canadian sites (Figure 3). Complementing this CONUS calibration, 47 basins are calibrated over Puerto Rico and 29 basins are calibrated across the Hawaii Islands, each using multi-year NAM-Nest-based forcing datasets unique to those domains, with observed precipitation supplied by the local RFC. The calibration and validation periods for Hawaii are, respectively, January 2005 to January 2010 and January 2010 to January 2014. Similarly, for Puerto Rico, calibration spans January 2009 to January 2014, with validation from January 2014 to September 2017.

Across the CONUS domain, a total of 14 model parameters (Table 4) are calibrated with an iterative Dynamically Dimensioned Search approach (Tolson & Shoemaker, 2007). A similar set of parameters, minus the MFSNO melt factor, are calibrated for the Hawaii and Puerto Rico domains. In this technique, the model is cycled over the calibration period 300 times to minimize an objective cost function. For the NWM, the objective cost function is a weighted Nash-Sutcliffe Coefficient of Efficiency (NSE) consisting of equal parts NSE (Nash & Sutcliffe, 1970) and NSE calculated for the log of the discharge (NSElog) using hourly streamflow observations. Once direct calibration is complete, parameters for the rest of the NWM domain are then derived via a regionalization process.

The regionalization process transfers optimized parameters from calibrated basins to the portions of the domain which cannot be directly calibrated due to a lack of observations or flow complications from water management. The basis for this procedure is the concept that basins featuring similar vegetation, topography, geology, soil properties and climate will have similar hydrologic parameters, and will be characterized by similar streamflow responses. To accomplish this regionalization, one of two techniques is applied, as determined through multi-year test simulations. For the full-physics NWM configuration (i.e. with the 250 m overland and subsurface routing functions active), the degree of similarity between each uncalibrated and calibrated basin is measured with the Gower's distance metric (Gower, 1971), while for the long-range NWM configuration an ecoregion-based approach is leveraged. Once matches are made between the directly calibrated donor basins and the uncalibrated receiver basins, parameters are mapped to the remainder of the NWM domain.

4 | NWM OUTPUT

Output from the NWM encompasses a wide range of hydrologic variables, which vary between modeling configurations based on end user needs and computational resource limitations. Real-time output totals nearly 1 Tb per day and is produced in a fully automated 24 × 7 operational fashion, the cadence of which varies by forecast configuration. By contrast, only one long-term retrospective run is released with each NWM version. Forced with AORC meteorological data, and executed without DA, the NWM v2.1 long-term retrospective run spans the period 1979–2020.

As extensive as the list of NWM output fields is, it is of very limited use unless it's paired with effective dissemination and visualization services. Towards this end, OWP has partnered with NCEP, NCAR, CUAHSI, and Big Data partners to leverage a set of foundational capabilities that will expand over time. The chief outlet for basic distribution of NWM data is the NOAA Operational Model Archive and Distribution System (NOMADS) hosted by NCEP. This publicly accessible system distributes data from models run operationally on the NOAA WCOSS, including the NWM. Full-domain NetCDF NWM output files are hosted on NOMADS and are made available for download via ftp and https protocols. Complementing this basic distribution are post-processed files, targeted for ingest into the NWS RFC operational forecast environment.

In addition to these operationally supported data sites, there exist a growing number of joint data ventures with both academic and private entities. These collaborations with Big Data partners and CUAHSI are especially important, given the challenges associated with distributing and processing the large amount of data produced by the NWM.

5 | MODEL PERFORMANCE

NWM simulation performance is assessed through a variety of approaches and metrics. Output from real-time forecast, historical reforecast, and retrospective analysis runs each provide a different window into NWM skill, and so are all assessed as part of the research to development to operations pipeline. These assessment activities document not only the skill of a particular version of a model but also the

TABLE 4 List of model parameters calibrated in NWM v2.1. Calibration scalar multiplier column indicates whether the parameter is adjusted uniformly within a calibration basin by a scalar multiplier during the parameter optimization process.

Relevant processes	Parameter	Description	Units	Default value	Range	Calibration scalar multiplier
Soil	*BEXP*	Pore size distribution index	Dimensionless	1	0.4–1.9	Yes
	SMCMAX	Saturation soil moisture content (i.e., porosity)	Volumetric fraction	1	0.8–1.2	Yes
	DKSAT	Saturated hydraulic conductivity	m/s	1	0.2–10	Yes
	RSURFEXP	Exponent in the resistance equation for soil evaporation	Dimensionless	5	1–6	No
Runoff	*REFKDT*	Surface runoff parameter	Unitless	1	0.1–4	No
	SLOPE	Coefficient to modify the drainage out the bottom of the last soil layer	0–1	0.3	0–1	No
	RETDEPRTFAC	Multiplier on retention depth limit	Unitless	1	0.1–20,000	No
	LKSATFAC	Multiplier on lateral hydraulic conductivity	Unitless	1000	10–10,000	No
Groundwater	*Zmax*	Maximum conceptual nonlinear reservoir depth	mm	50	10–250	No
	Expon	Exponent in the conceptual nonlinear reservoir module controlling rate of drainage as a function of depth	Dimensionless	3	1–8	No
Vegetation	*CWPVT*	Canopy wind parameter for canopy wind profile formulation	1/m	1	0.5–2	Yes
	VCMX25	Maximum carboxylation at 25°C	$\mu mol/m^2/s$	1	0.6–1.4	Yes
	MP	Slope of Ball-Berry conductance relationship	Unitless	1	0.6–1.4	Yes
Snow	*MFSNO*	Melt factor for snow depletion curve	Dimensionless	1	0.25–2	Yes

TABLE 5 Overview of major new features implemented within each version of the NWM.

Operational NWM version	Notable new features
V1.0 (2016)	Initial version of the NWM with CONUS coverage
V1.1 (2017)	Expansion along CONUS border, first detailed calibration of parameters, increased forecast cycling frequency and length
V1.2 (2018)	Improved calibration and streamflow data assimilation, additional domain expansion along CONUS border
V2.0 (2019)	Expansion to Hawaii, seven-member medium-range CONUS ensemble, use of RFC MPE precipitation, improved calibration, increased reservoir representation, improved forcing downscaling and model physics
V2.1 (2021)	Expansion to Puerto Rico/US Virgin Islands and Great Lakes Drainage domains, implementation of reservoir outflow data assimilation, forcing upgrades, new open-loop configurations, improved physics and calibration
V3.0 (2023)	Total water level coastal coupling component, expansion to Alaska, ingestion of National Blend Models forcing, improved runoff module, and improved physics and calibration

version-over-version evolution as new features are added (Table 5). Together they help focus development efforts on where they are most needed, and inform forecasters as to the level of confidence which characterizes NWM forecasts.

5.1 | Streamflow verification—Retrospective output

Assessment of NWM streamflow is based on comparison against USGS hourly streamflow observations. As with verification of meteorological models, the number of model output points (over 2.7 million) dwarfs the number of available observation sites (~10,000). However, the sequentially connected nature of many of the stream reaches acts as a mitigating factor, allowing downstream gauges to be used to verify, by proxy, nearby upstream stretches of the river. This effectively reduces the spatial gap in verification data, but does not address another verification challenge—anthropogenic influences. Given the profound impact that dams, diversions and other similar elements have on natural streamflow, USGS data are divided up into two NWM verification sets: (1) The full USGS observation collection of approximately 10,000 stations, and (2) A reference stream gauge dataset from the Geospatial Attributes of Gages for Evaluating Streamflow version II (GAGES-II; Falcone, 2017) database, which includes a subset of non- or lightly-regulated streams (approximately 1600 stations).

Using the full USGS dataset as the baseline, Figure 4 displays NWM hourly percent bias values as CONUS-aggregate histograms. Output is from a continuous retrospective simulation using the full-routing formulation of the model. This validation run was executed from 2013 to 2016, directly after the 2008 to 2013 NWM calibration period, and separate from the long-term 1979–2020 retrospective run. It was driven with AORC meteorological forcing data, and did not assimilate USGS streamflow observations. Progressing from NWM V1.0 through V2.1, each new version of the NWM has seen a related reduction in simulated streamflow bias. In particular, while 31% of verification sites have a bias of ±20% (the most favorable category) in V1.0, 52% of those same sites reach that level of accuracy in V2.1. This improvement is accompanied by a favorable tightening of the histogram distribution, with a decrease in wet and dry extremes. The spatial distribution of the errors are somewhat uneven, with wet biases more prevalent in the middle of the CONUS than in other regions. Removing the influence of streamflow regulation from the assessment, Figure 5 displays bias statistics from the same simulation, but this time validated against the GAGES-II subset of observations. The same general findings hold, although performance levels are predictably higher given the lack of complications introduced by reservoirs and diversions. Bias levels of ±20% are seen in 36% of validation locations in V1.0, increasing to 68% by V2.1 of the NWM.

An assessment of hourly model correlation and normalized Nash Sutcliffe Efficiency (NNSE, Nossent & Bauwens, 2012) provides additional insight into NWM performance. Verification results at GAGES-II locations are presented in Figures 6 and 7 using an approach similar to that taken with bias. As with bias, a strong improvement in results is seen version-over-version. While only 18% of verified reaches have a correlation of 0.8 or greater in V1.0, improvements between versions leads to a value of 52% by V2.1 (Figure 6). Similarly, only 43% of verification basins have an hourly NNSE of 0.6 or greater in v1.1, rising to 81% in v2.1. The spatial variation in results for both metrics is also somewhat similar to that of bias, with lower values noted over the midsection of the CONUS.

Verification of the simplified NWM configuration used in Long-Range CONUS applications (i.e. without the 250 m terrain routing) was also conducted for GAGES-II locations over this same retrospective time period. Summarizing results from V1.2 through V2.1, Figure 8 shows the improvements in streamflow correlation between model versions. Overall, model results are less favorable than those seen in Figure 7, which is to be expected given the lack of explicit surface and sub-surface routing in this simplified, long-range configuration of the NWM.

Complementing the CONUS verification is a parallel assessment covering the NWM domains outside of the CONUS, namely Hawaii and Puerto Rico. Over the relatively new Hawaii domain, great strides in performance were made between V2.0—when the domain was first

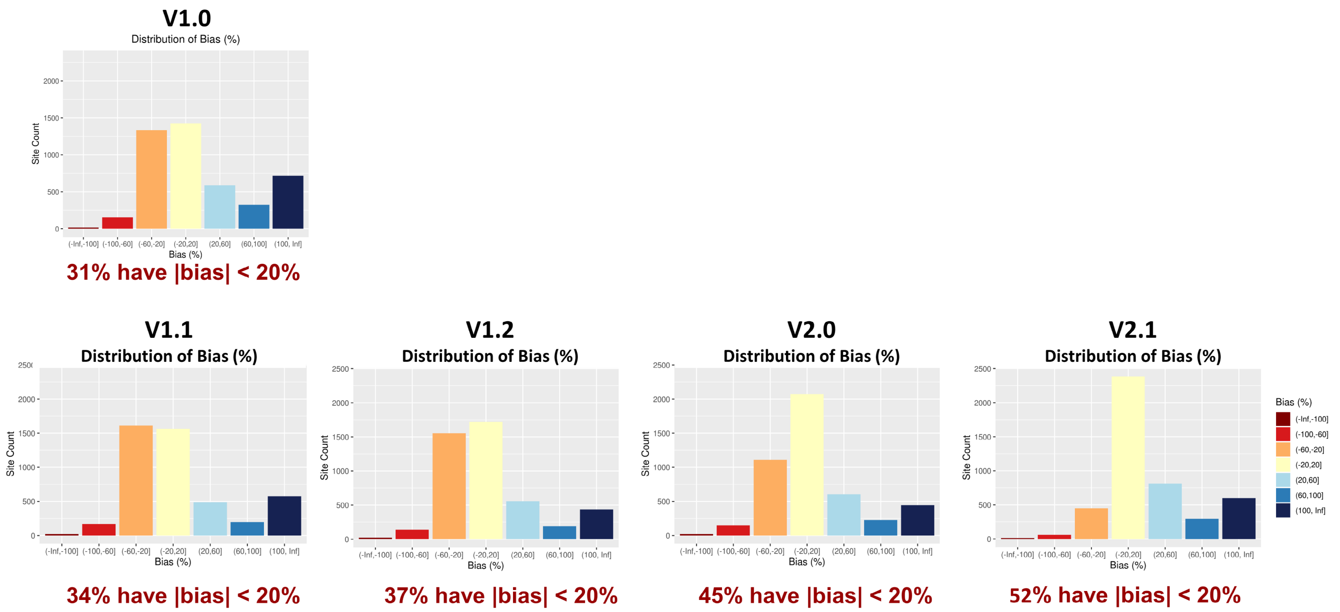


FIGURE 4 Hourly NWM streamflow bias (%) from version 1.0 through version 2.1. Statistics are for the 2013–2016 post-calibration validation period are based on the full set of CONUS USGS gauges.

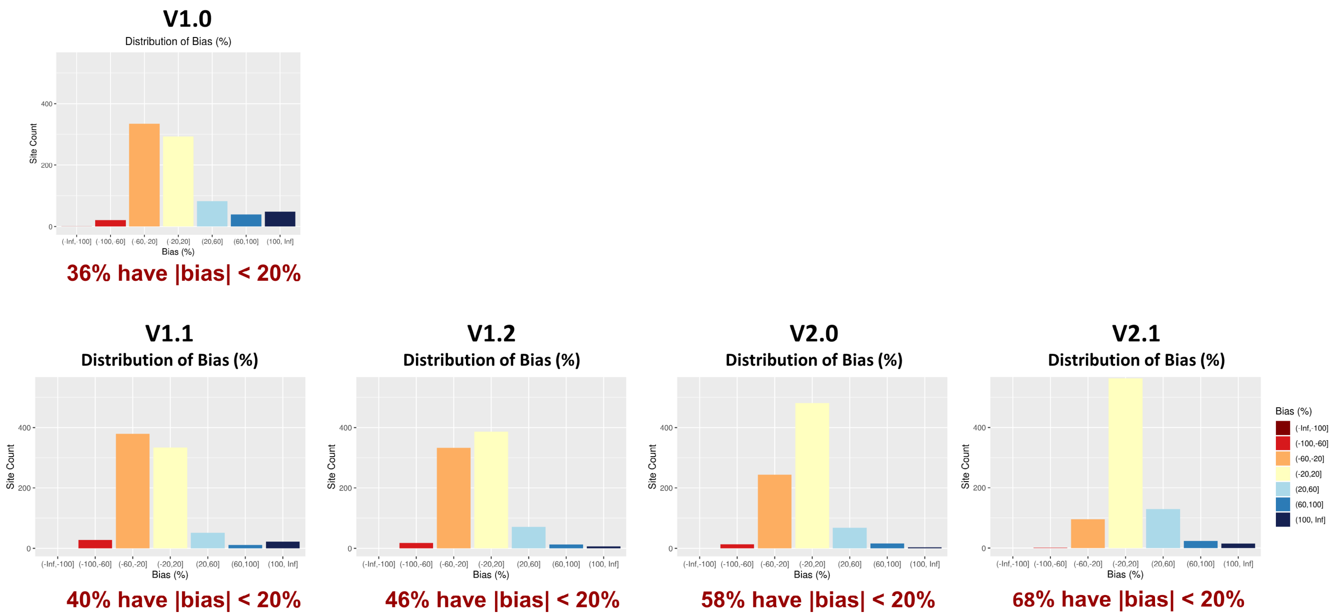


FIGURE 5 Hourly NWM streamflow bias (%) from version 1.0 through version 2.1. Statistics are for the 2013–2016 post-calibration validation period are based on the Gauges II subset of CONUS USGS gauges.

introduced—and v2.1. Figure 9 depicts a significant reduction in the model's overall wet bias, with biases remaining high in areas with significant anthropogenic impacts (e.g. diversions, urbanization) and particularly dry regions. An increase in correlation (not depicted) is noted as well. Across the NWM's other island domain, which includes Puerto Rico and the USVI, strong initial performance metrics are noted in v2.1, the version in which it was implemented. Here, 51% of the verification sites show a streamflow bias of less than ±20%, while 48% (79%) of sites display an hourly (daily) streamflow correlation greater than 0.6 (Figure 10).

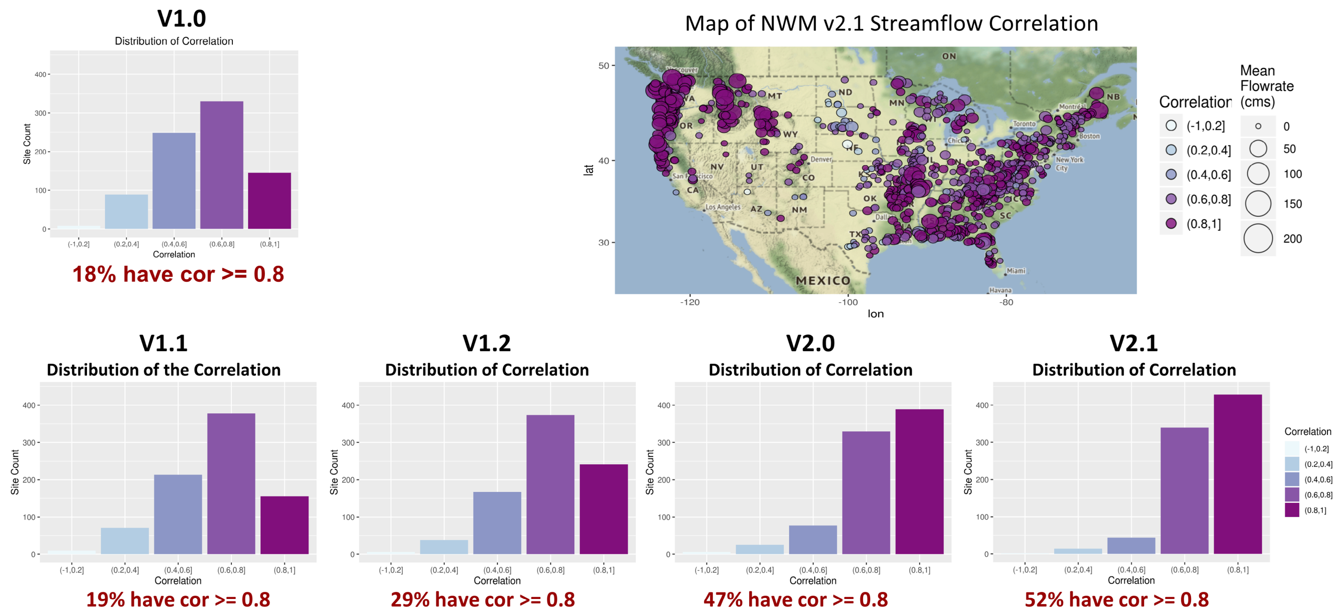


FIGURE 6 Hourly NWM streamflow correlation from version 1.0 through version 2.1. Statistics are for the 2013–2016 post-calibration validation period are based on the Gauges II subset of CONUS USGS gauges.

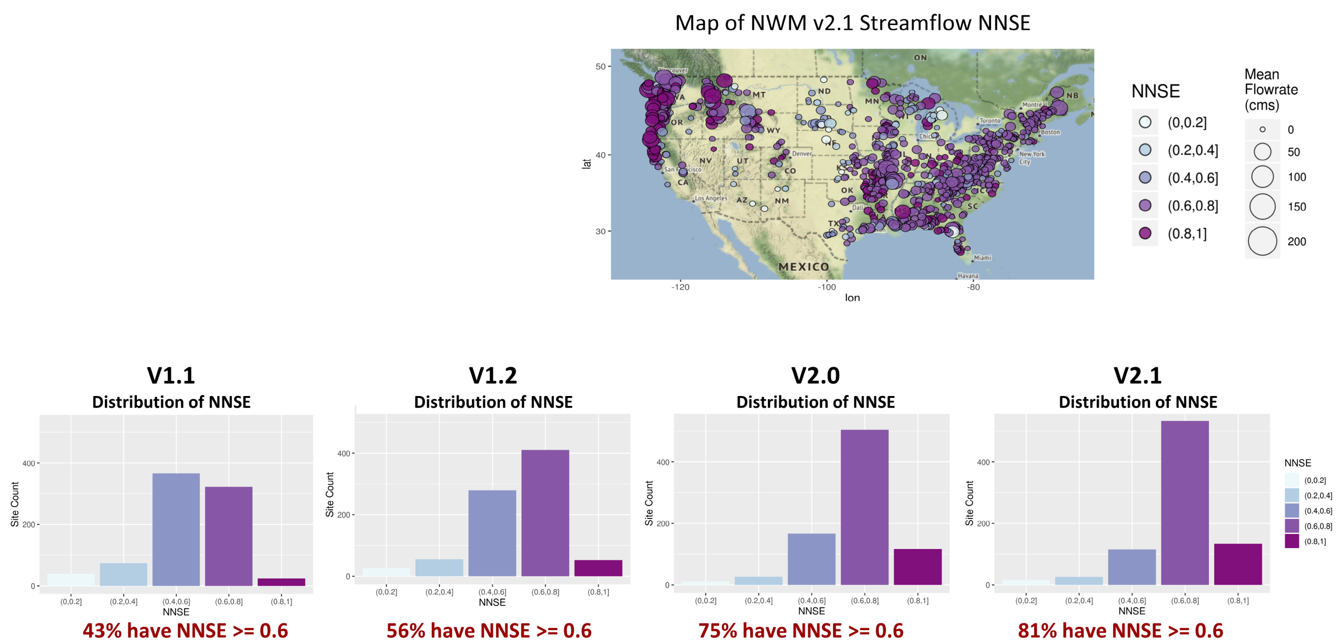


FIGURE 7 Hourly NWM streamflow Normalized Nash Sutcliffe Efficiency (NNSE) from version 1.1 through version 2.1. Statistics are for the 2013–2016 post-calibration validation period are based on the Gauges II subset of CONUS USGS gauges.

5.2 | Snowpack verification

The NWM contains a multi-layered, energy-balance-based snowpack formulation which tracks snowpack conditions on the 1-km LSM grid over the entire model domain. The performance of the snowpack representation in each model version has been evaluated against Natural Resources Conservation Service (NRCS) SNOwpack TELemetry (SNOTEL) station observations and against the NOAA OWP gridded Snow Data Assimilation System (SNODAS) analysis product (Carroll et al., 2006). Figure 11 shows a comparison of model performance between v2.0 and v2.1 of the NWM against these observations for retrospective run spanning 2011–2017. Figure 11a shows that, in general, a modest bias

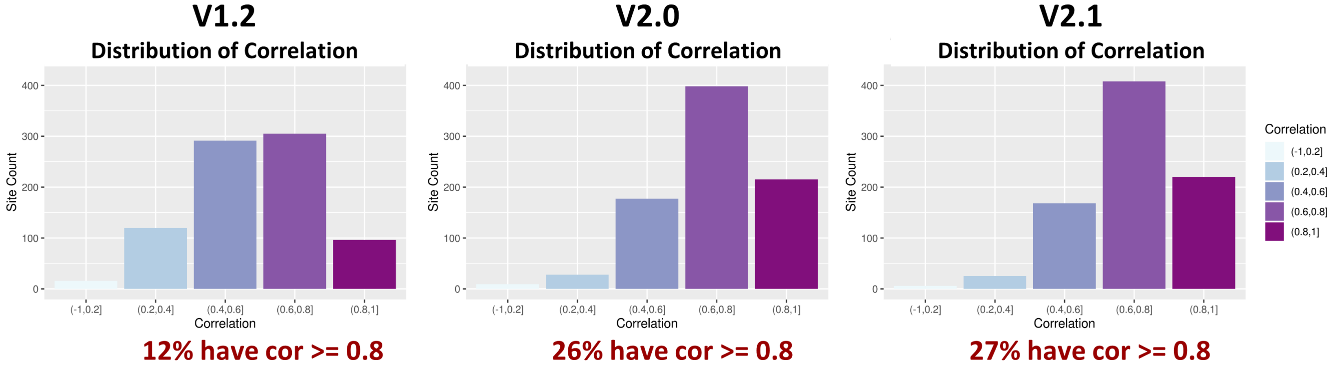


FIGURE 8 Hourly NWM streamflow correlation from version 1.2 through version 2.1. Statistics are for the 2013–2016 post-calibration validation period are based on the Gauges II subset of CONUS USGS gauges and are computed using the simplified Long-Range configuration of the NWM.

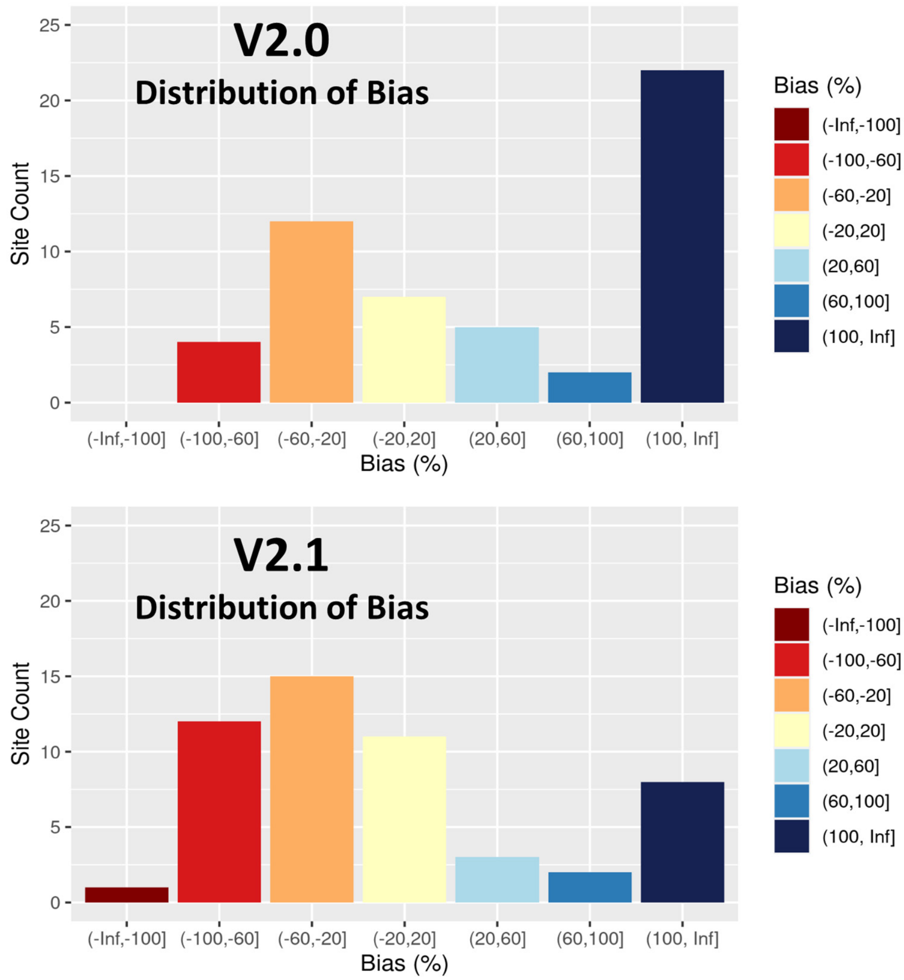


FIGURE 9 Hourly NWM streamflow bias (%) for versions 2.0 and 2.1. Statistics are for the 2010–2013 post-calibration validation period are based on the full set of Hawaii USGS gauges.

improvement in modeled snowpack performance—as quantified by local model minus observed values and aggregated across RFC regional SNOTEL locations—was achieved in v2.1 for most, though not all, regions. Figure 11b shows the Central Rocky Mountain, regionally averaged time-series of snow water equivalent for the two NWM versions and the SNODAS product. Modest improvement in v2.1 snowpack accumulation volume is found both in terms of the regional mean and max/min range of values during the 2011–2017 retrospective assessment period.

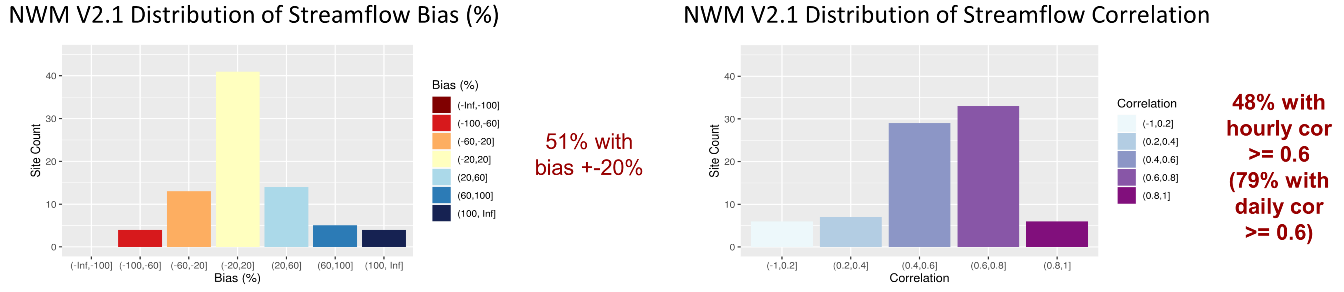


FIGURE 10 Hourly NWM streamflow bias (%) and correlation for version 2.1. Statistics are for a 2009–2017 evaluation period are based on the full set of Puerto Rico USGS gauges.

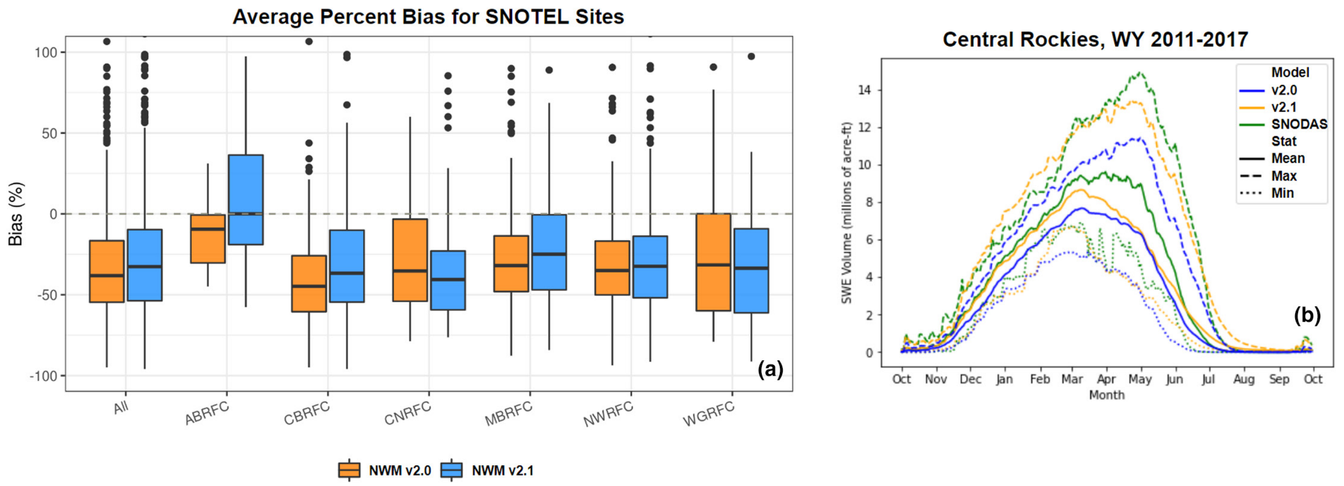


FIGURE 11 Assessment of NWM v2.0 and v2.1 snowpack against SNOwpack TELemetry (SNOTEL) (percent bias, left; a) and Snow Data Assimilation System (SWE volume, right; b) data for 2011–2017 retrospective assessment period. In (b), line colors denote model, line styles denote statistic.

Improvements in snowpack conditions are attributed to improvements in the snow physics formulation related to snowpack storage and release of melt water as well as to improved calibration of some of the snowpack model parameters during automated calibration.

5.3 | Streamflow verification—Real-time operational output

Of equal importance with the preceding suite of verification statistics derived from the assessment of NWM retrospective validation simulations is a parallel set of statistics based on real-time operational NWM forecast output. These assessments yield information which then guides use of the model by forecasters and other end users. Unlike the historical runs which are forced with observed precipitation, real-time NWM forecasts are forced with NWP-based precipitation. This changes the error profile of the model significantly, with input precipitation now an increased source of error.

Focusing on the short-range configuration, Figure 12 depicts the improvements in performance between NWM versions 2.0 and 2.1. The sizable overestimation of peak discharge in v2.0 (especially at longer lead times) is greatly reduced in v2.1, with the distribution of the median absolute percent peak bias centered near zero. This is mirrored by improvements in event-type statistics including probability of detection (POD), false alarm ratio (FAR) and critical success index (CSI). Peak streamflow timing error is similar in v2.1 and v2.0, with the forecast streamflow too quick to peak.

As highlighted in Figure 13, overall improvement is also noted in the medium-range real-time forecast configuration. Peak bias and peak timing errors improve during the first 6 h, but change little in the hours that follow. Model improvement is more pronounced when examining categorical flood prediction scores. Here, notable decreases are seen in the FAR, accompanied by sizable increases in the CSI. Similar results are found for days 4–10, with the medium-range ensemble means exhibiting higher scores (not shown). Focusing on the performance of

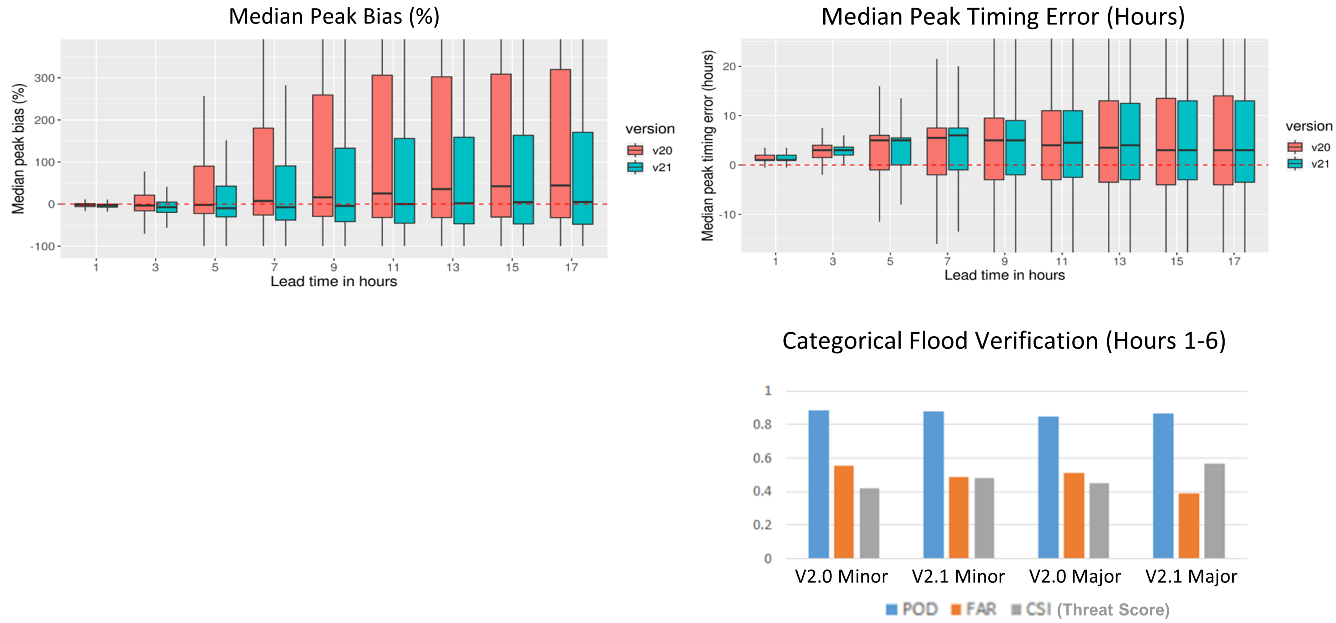


FIGURE 12 Median peak bias (%, top left), median peak timing error (hours, top right), and categorical flood verification metrics (bottom right) for NWM v2.1 versus NWM v2.0, for NWM short-range configuration. Calculated with data from July 1 to August 9, 2020.

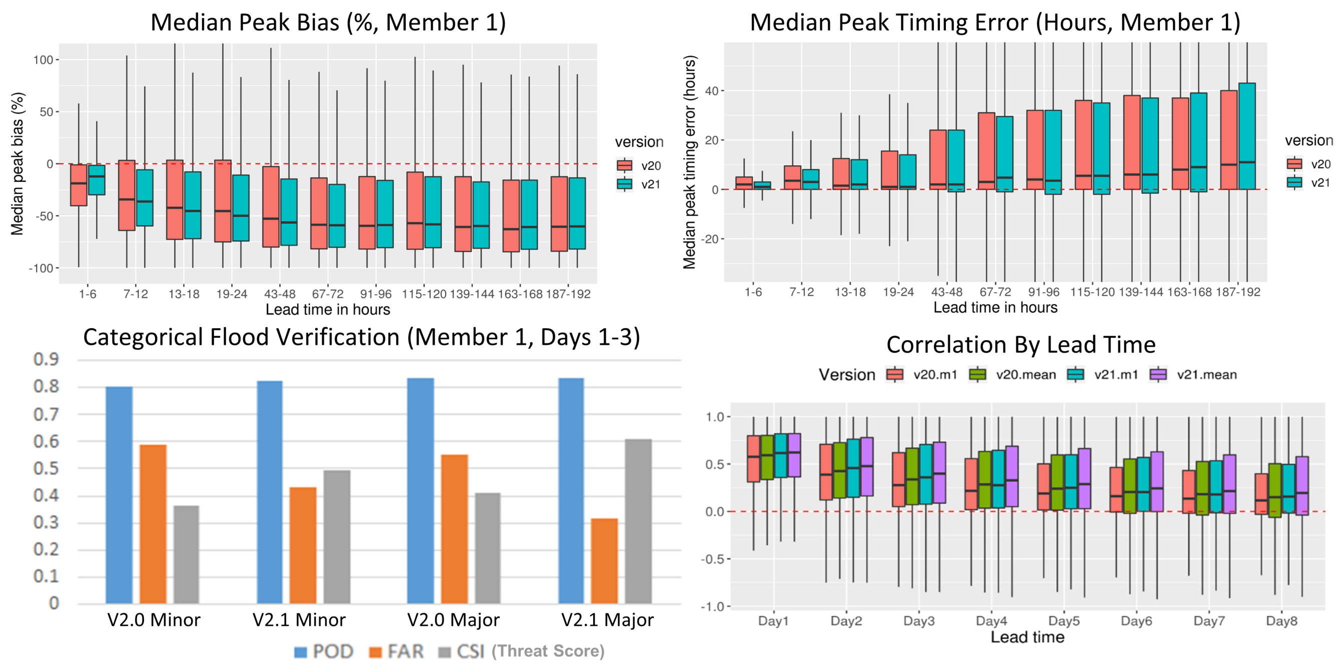


FIGURE 13 Median peak bias (%, top left), median peak timing error (h, top right), categorical flood verification metrics (bottom left) for member 1 of the NWM medium-range configuration, with categorical scores calculated for approximately 3000 sites. Additionally, correlation by lead time member 1 and the ensemble mean for both NWM v2.0 and v2.1 (bottom right). Calculated with data from July 1 to August 9, 2020.

streamflow prediction downstream from reservoirs, Figure 14 isolates the impact of upgrades made to the reservoir module. In these two representative examples, the application of observation persistence and the ingestion of RFC-sourced reservoir forecasts in NWM v2.1 lead to sizable benefits.

Switching focus to the NWM's offshore domains, large version-over-version improvements in Hawaii streamflow prediction skill are shown in Figure 15. The large overestimation of peak discharge in v2.0 is significantly reduced in v2.1 across all lead times, and the peak timing error

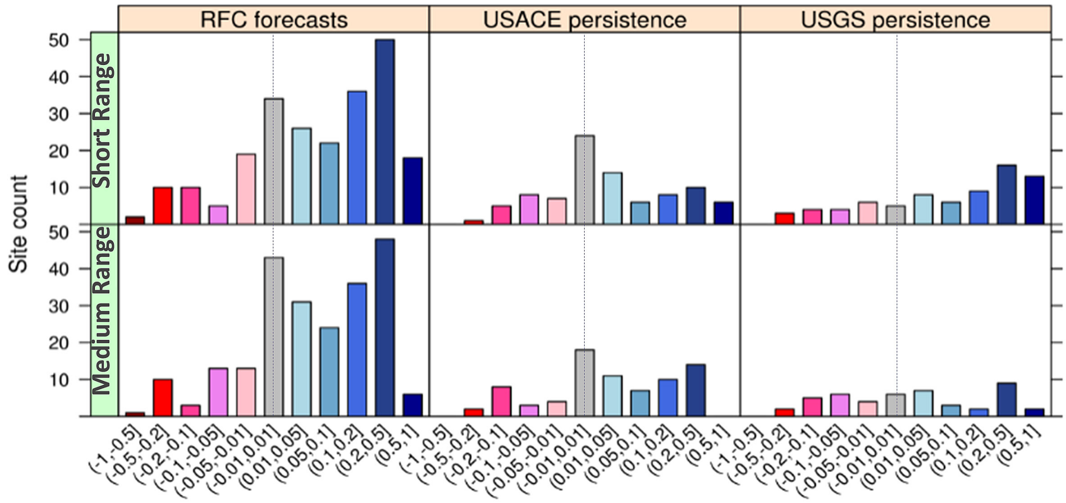
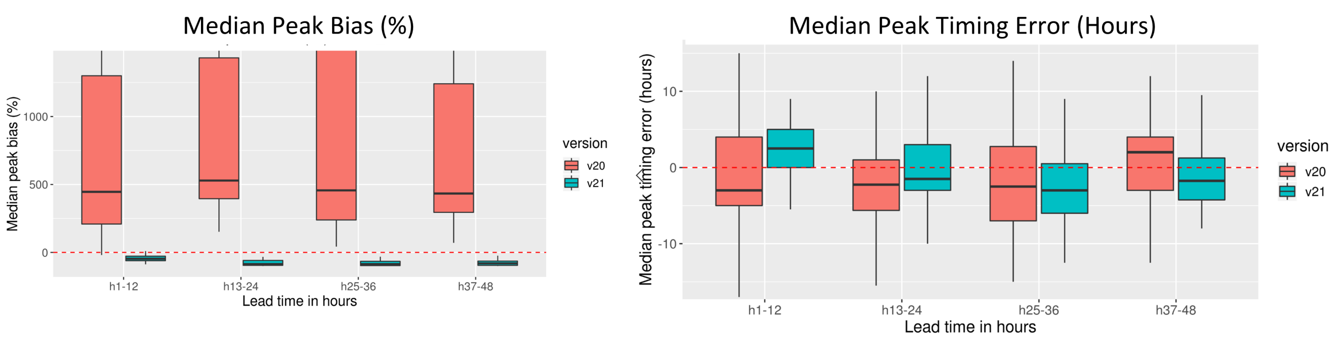


FIGURE 14 Relative change in NNSE from NWM v2.0 to v2.1 at sites with three different sources of reservoir outflows for both the short-range (top row) and medium-range (bottom row) forecast configurations. Note that the plots show the version-on-version change in NNSE, not the actual NNSE values, and are calculated with data from July 1 to August 9, 2020.



Categorical Flood Verification (Hours 1-12)

Observed Flood Events	**0**
V2.0 False Alarm Events	**21 Minor, 2 Major**
V2.1 False Alarm Events	**0**

FIGURE 15 Event-based assessments by lead time for NWM v2.0 and v2.1 Short-Range forecasts over Hawaii domain. Median peak bias (%, left), median peak timing error (hours, upper right), and table of false alarms (bottom right). Calculated with data from July 1 to August 9, 2020.

distribution has a smaller range in v2.1 than v2.0. Additionally, categorical flood detection is greatly improved in V2.1 (similar results for hours 13–48, not shown), with a reduction from 23 false alarms in V2.0 to zero in v2.1. Improvements in the selection and processing of source precipitation data as well as improvements in calibration methods are attributed to this increase in skill over Hawaii.

The NWM's second domain outside of the CONUS—Puerto Rico—is new for v2.1, and so there is no prior NWM version against which to compare results. However, it is still instructive to examine the performance of this first implementation. Figure 16 shows that peak discharge is generally underestimated across all lead times, while the median peak timing error is relatively small (median within ± 3 h). Assessing the output from a categorical event perspective, it can be seen that this first implementation of the Puerto Rico domain displays reasonable categorical flood forecast skill, with a POD of 0.7 and a FAR of 0.42. These scores support initial use of the model, complementing information streams that are already available, and future versions of the NWM will improve upon these scores.

An in-depth comparison of NWM and RFC-produced streamflow forecasts has the potential to increase understanding of the relative merits of each approach and aid in the effective use of each source of data. Such a study would provide insight into the use of a complementary mix

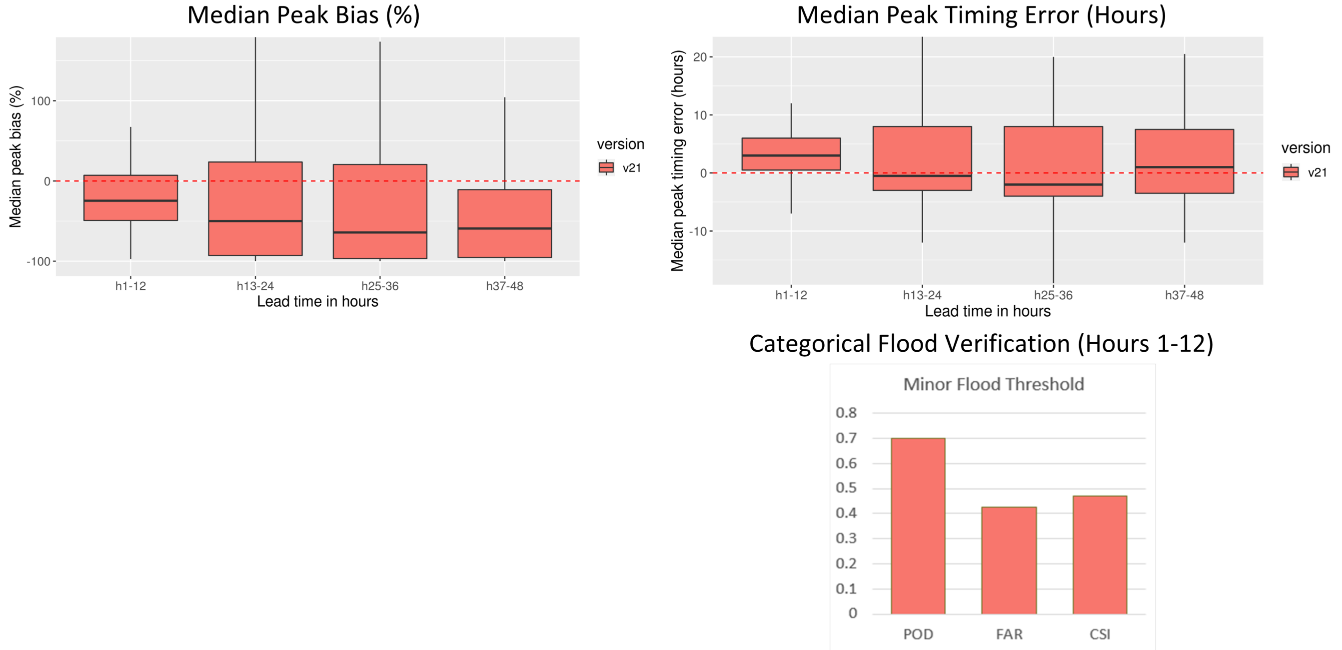


FIGURE 16 Event-based assessments by lead time for NWM v2.1 Short-Range Forecast over Puerto Rico domain. Median peak bias (%, left), median peak timing error (hours, upper right), and table of false alarms (bottom right). Calculated with data from July 1 to August 9, 2020.

of data, across various hydrologic situations in time and space. However, a meaningful intercomparison is hampered by the many differences between NWM and RFC forecast systems, as well as from RFC-to-RFC. These increase the complexity of an assessment and interpretation of the results to such an extent that it moves beyond the scope of this overview article. In particular, RFC forecasts are produced over differing spatial regions, at a different cadence, using differing input forcing data sources and with forecaster modifications to forcing inputs, model states and parameters. A direct and informational comparison against the NWM would entail conducting reforecasts or reanalyses with both platforms using the same forcing data, but would still fail to capture important differences in model performance for the aforementioned reasons. Until completion of a study focusing exclusively on such an in-depth and multifaceted assessment, the summary statistics in this NWM overview article serve to highlight the current performance characteristics of the NWM, and provide both a basis for use and a foundation for a future, in-depth intercomparison article.

6 | ONGOING DEVELOPMENT

6.1 | Operational enhancements

Established in 2016, the NWM is a relatively young modeling system. The NWM continues to undergo cyclic upgrades so as to improve and expand version-over-version capabilities guided by the assessments described above. Certain foundational components of the NWM are improved with each upgrade, while other components are added for the first time, or may only receive sporadic updates. Improvements which accompany each new model version include updated calibrated parameters, fixes to the hydro-fabric connections and waterbody attributes, along with updates to the workflow, model and MFE to improve execution robustness and efficiency.

Nowhere is the drive towards improved representation of hydrologic processes more important than along coastal regions. In this area, over 100 million people currently lack a forecast of the integrated impacts of freshwater, storm surge, waves and tidal flooding. Within version 3.0 of the NWM, a new routing capability will support linkage of NWM freshwater modeling capabilities to a coastal-estuary model, the Semi-implicit Cross-scale Hydroscience Integrated System Model (SCHISM) (Zhang, Ye, et al., 2016). Atmospheric forcing will be drawn from the existing set of NWM forcing data described above, while ocean forcing will be drawn from the Surge and Tide Operational Forecast System (STOFS; Funakoshi et al., 2012) and Probabilistic Tropical Storm Surge (P-SURGE; Taylor & Glahn, 2008) models. This approach will be applied over the East, Gulf and Pacific coasts of the CONUS, along with the coastlines of the Hawaii, Puerto Rico, and US Virgin Islands, within the AnA, Short-Range, and Medium-Range forecast configurations. It will provide enhanced guidance to emergency responders and will improve the accuracy of NWM-based flood inundation maps along the coast.

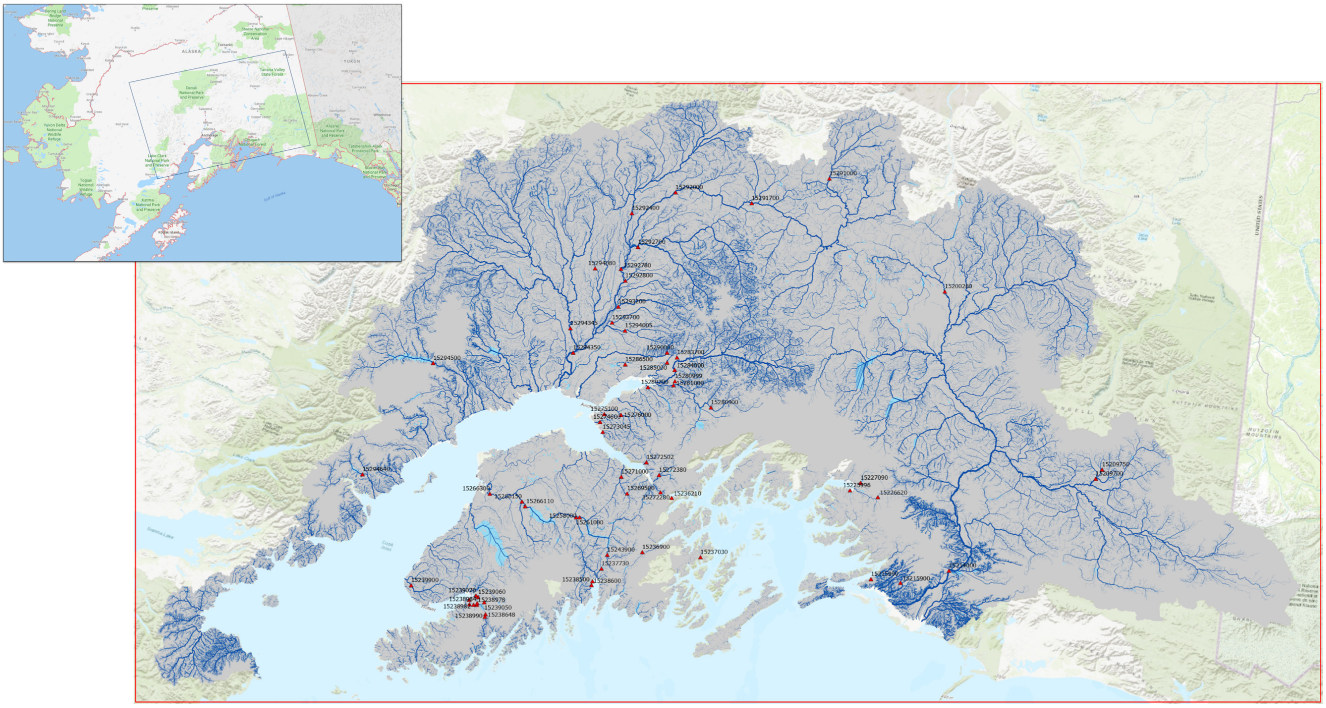


FIGURE 17 Depiction of the new NWM Alaska domain, to be released in version 3.0 of the model.

Service provision will be further improved via first-time NWM coverage for Alaska, wherein an NWM configuration is scheduled for operational deployment over the Cook Inlet and Copper River Basin regions (Figure 17). Several upgrades were introduced into the NWM in order for it to better function in this cold land process-dominated area. Chief among these are a linkage with a snow-ice model (Vionnet et al., 2012) and the capability for ingest of RFC-provided forecasts of streamflow from glacial lake outbursts. Combined with forcing and hydrofabric datasets tailored to the needs of this particular domain, this implementation will bring much-needed comprehensive and complementary hydrologic guidance to the south-central Alaska region.

Complementing the foundational enhancements above will be a shift to updated land cover and channel datasets, improved treatment of infiltration, the creation of a dynamic parameter update capability, and use of the National Blend of Models as an input forcing source.

Taken together, the wide-ranging process modules and forcing enhancements discussed above combine with the NWM's underlying distributed modeling structure to support key nationwide hydrologic applications in a way not before possible, and connect back to the driving factors behind the NWM's creation. A prime example of this is the advent of coast-to-coast flood inundation mapping (Aristizabal et al., 2023). Critical for emergency managers in times of extreme flow, this application leverages NWM analyses and forecasts to provide distributed inundation information down to the neighborhood structure level. The NWM's fine scale NHD-based stream network, along with its representation of multiple processes including coastal total water level, supports seamless production of this information from summit to sea, and will continue to evolve with each iteration of the NWM.

6.2 | Community assessment and development

To date a number of published works have documented various performance aspects of the NWM. These works have analyzed several aspects of the skill of reservoir inflow forecast performance (Viterbo, Read, et al., 2020) and snow model performance (Garousi-Nejad & Tarboton, 2022). Viterbo, Mahoney, et al. (2020) explored the ability of the NWM to forecast specific high-impact flooding events while Tijerina et al. (2021) provided a multi-metric assessment of NWM model performance in comparison with another state-of-the-art hydrologic model. The research community surrounding the NWM has also produced a number of other papers centering on development aspects for specific processes within the NWM, including the improvement of the NWM for semi-arid environments (Lahmers et al., 2019, 2022), the development of hydro-geo-fabric data for the Great Lakes region (Mason et al., 2019), the assimilation of remotely sensed vegetation data (Elmer et al., 2022), the enhancement of channel routing physics (Read et al., 2023) and the impact of tile drainage processes on NWM simulation performance (Valayamkunnath et al., 2022).

Now at version 2.1, with NWM v3.0 arriving in late 2023, the continued success of the NWM ultimately depends on continual improvement of model accuracy and capabilities—something which, in turn, depends on a vibrant and robust development community with a clear and efficient link to the NWS operational pipeline. With this in mind, OWP is embarking upon a major, multi-year effort, to build on and expand the NWM development and applications community. OWP and agency partners are improving upon the existing modularity of NWM processes and are establishing the Next Generation Water Resources Modeling Framework (NextGen) (Ogden et al., 2021). NextGen is a standards-based model interoperability framework that allows for the use of spatially varying modeling techniques which are the most appropriate for each region. Utilizing the Basic Model Interface (Hutton et al., 2020) coupling standard, NextGen facilitates linkage to a wide range of community-sourced modules. The NextGen framework also employs the National Hydrologic Geospatial Fabric Reference Hydrofabric (Bock et al., 2022) developed jointly by USGS and NOAA. This reference hydrofabric adopts the WaterML version 2.0 Hy_Features data model (OGC, 2017). The overarching end goal for all of these activities is a system which supports community development and funnels innovation into a common standards-based platform that can be leveraged for both research and operations across a wide range of scales and applications. The NOAA-NWS OWP anticipates that Version 4.0 of the NWM will use the NextGen Framework beginning in late 2025 or 2026.

7 | CONCLUDING THOUGHTS

With the NWM in NWS operations since 2016, the hydrologic community has a nationwide platform which will further both operational and research interests. The NWM features forecast horizons from 18 h to 30 days, covers the CONUS and nearby regions along with Hawaii, Puerto Rico and US Virgin Islands, and soon south-central Alaska. It provides hydrologic output for over 2.7 million stream reaches and land surface output on high-resolution 100 m to 1 km grids. The model continues to evolve through version-over-version upgrades, providing powerful guidance to complement the vital forecasts already being produced by the RFCs, and filling in spatial gaps where little to no hydrologic information was previously available—a critical advance, as population growth combined with aging infrastructure increases hydrologic vulnerability. The model's density of coverage ensures a gap filling, quantifiable increase in the proximity of guidance, with a current CONUS domain-wide average distance to an NWM output location of less than 975 meters. This supports access not only to relevant NWM streamflow and related hydrologic guidance, but to key advances in seamless flood inundation mapping products.

Beyond increases in coverage, the NWM's unique mix of high spatial resolution, multiple forecast horizons, nationwide domain, operational robustness and accessible code allow it to serve as focal point for government agencies across the water spectrum, from NOAA to the USGS, USACE, EPA, USDA, FEMA, NASA, and the USBR. At the same time, increasing links with academia and private industry are providing the foundation for a rich and responsive development environment. Combined with Big Data partnerships and strengthened use of GIS-based services to lower data access barriers, the NWM provides information needed to further the NWS mission of protecting lives and property, and to support a broad range of other hydrologic applications in ways not before possible.

AUTHOR CONTRIBUTIONS

Brian Cosgrove: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing – original draft; writing – review and editing. **David Gochis:** Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing – original draft; writing – review and editing. **Trey Flowers:** Conceptualization; funding acquisition; methodology; project administration; resources; supervision; writing – review and editing. **Aubrey Dugger:** Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing – review and editing. **Fred Ogden:** Conceptualization; formal analysis; investigation; methodology; writing – review and editing. **Tom Graziano:** Conceptualization; funding acquisition; methodology; project administration; resources; supervision; writing – review and editing. **Ed Clark:** Conceptualization; funding acquisition; methodology; project administration; resources; supervision; writing – review and editing. **Ryan Cabell:** Conceptualization; data curation; investigation; methodology; software; validation; visualization. **Nick Casiday:** Funding acquisition; project administration; resources; supervision. **Zhengtao Cui:** Conceptualization; data curation; investigation; methodology; software; validation; visualization. **Kelley Eicher:** Conceptualization; data curation; formal analysis; software; validation; visualization. **Greg Fall:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Xia Feng:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization; writing – review and editing. **Katelyn Fitzgerald:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Nels Frazier:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Camaron George:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Rich Gibbs:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Liliana Hernandez:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Donald Johnson:** Conceptualization; data curation; software. **Ryan Jones:** Funding acquisition; project administration; resources; supervision. **Logan Karsten:** Conceptualization; data curation; formal analysis; investigation;

methodology; software; validation; visualization. **Henok Kefelegn:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **David Kitzmiller:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Haksu Lee:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Yuqiong Liu:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Hassan Mashriqui:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **David Mattern:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Alyssa McCluskey:** Funding acquisition; project administration; resources; supervision. **James McCreight:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Rachel McDaniel:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization; writing – review and editing. **Alemayehu Midekisa:** Conceptualization; data curation; investigation; methodology; software; validation; visualization. **Linlin Pan:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Andy Newman:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Cham Pham:** Conceptualization; data curation; investigation; methodology; software; validation; visualization. **Arezoo RafieeiNasab:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Roy Rasmussen:** Conceptualization; funding acquisition; methodology; project administration; resources; supervision. **Laura Read:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Mehdi Rezaeianzadeh:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Fernando Salas:** Conceptualization; data curation; formal analysis; investigation; methodology; software; supervision; validation; visualization. **Dina Sang:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Kevin Sampson:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization; writing – original draft. **Tim Schneider:** Conceptualization; funding acquisition; methodology; project administration; resources; supervision. **Qi Shi:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Gautam Sood:** Data curation; software. **Andy Wood:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization; writing – original draft. **Wanru Wu:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **David Yates:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization. **Wei Yu:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation. **Yongxin Zhang:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

FUNDING INFORMATION

None.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available via the HydroShare data repository at https://doi.org/10.4211/hs.f5041b451e66468c962095f33683738a.

ORCID

Brian Cosgrove https://orcid.org/0000-0002-3921-0553
Greg Fall https://orcid.org/0000-0002-5059-4703
Liliana Hernandez https://orcid.org/0000-0002-5259-0968
Kevin Sampson https://orcid.org/0000-0002-5537-7775
Qi Shi https://orcid.org/0000-0002-5338-8112

REFERENCES

Aristizabal, F., F. Salas, G. Petrochenkov, T. Grout, B. Avant, B. Bates, R. Spies, N. Chadwick, Z. Wills, and J. Judge. 2023. “Extending Height above Nearest Drainage to Model Multiple Fluvial Sources in Flood Inundation Mapping Applications for the U.S.” *National Water Model. Water Resources Research* 59: e2022WR032039. https://doi.org/10.1029/2022WR032039.

Arnault, J., B. Fersch, T. Rummler, Z. Zhang, G.M. Quenum, J. Wei, M. Graf, and H. Kunstmann. 2021. “Lateral Terrestrial Water Flow Contribution to Summer Precipitation at Continental Scale—A Comparison between Europe and West Africa with WRF-Hydro-Tag Ensembles.” *Hydrological Processes* 35: e14183. https://doi.org/10.1002/hyp.14183.

Ashley, S.T., and W.S. Ashley. 2008. “Flood Fatalities in the United States.” *Journal of Applied Meteorology and Climatology* 47(3): 805–18.

Benjamin, S.G., S.S. Weygandt, J.M. Brown, M. Hu, C.R. Alexander, T.G. Smirnova, J.B. Olson, et al. 2016. “A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh.” *Monthly Weather Review* 144(4): 1669–94.

Bock, A.R., D.L. Blodgett, J.M. Johnson, M. Santiago, and M.E. Wieczorek. 2022. "PROVISIONAL: National Hydrologic Geospatial Fabric Reference and Derived Hydrofabrics." U.S. Geological Survey Data Release. https://doi.org/10.5066/P9NFPB5S.

Carley, J.R., E. Rogers, B.S. Ferrier, E. Aligo, Z. Janjic, S. Liu, Y. Lin, et al. 2017. "The Nested Domains of Version 4 of the NAM Forecast System." 28th Conf. on Weather Analysis and Forecasting/24th Conf. on Numerical Weather Prediction, Seattle, WA, Amer. Meteor. Soc., 1204. https://ams.confex.com/ams/97Annual/webprogram/Paper311148.html.

Carroll, T.R., D.W. Cline, C. Olheiser, A. Ros, A. Nilsson, G. Fall, C. Bovitz, and L. Li. 2006. "NOAA's National Snow Analyses." Proceedings of the 74th Annual Meeting of the Western Snow Conference 74: 13.

Chow, V.T. 1959. *Open Channel Hydraulics*. New York: McGraw-Hill.

Coordinated Ocean Observations and Research Act of 2020. 2020. "Public Law No: 116–271." https://www.congress.gov/bill/116th-congress/senate-bill/914/text.

Cosgrove, B.A., D. Lohmann, K.E. Mitchell, P.R. Houser, E.F. Wood, J.C. Schaake, A. Robock, et al. 2003. "Real-Time and Retrospective Forcing in the North American Land Data Assimilation System (NLDAS) Project." *Journal of Geophysical Research: Atmospheres* 108(22): GCP 3-1–3-12. https://doi.org/10.1029/2002jd003118.

Dowell, D.C., C.R. Alexander, E.P. James, S.S. Weygandt, S.G. Benjamin, G.S. Manikin, B.T. Blake, et al. 2022. "The High-Resolution Rapid Refresh (HRRR): An Hourly Updating Convection-Allowing Forecast Model. Part I: Motivation and System Description." *Weather and Forecasting* 37(8): 1371–95.

Elmer, N.J., J.R. Mecikalski, A.L. Molthan, D.J. Gochis, and U. Nair. 2022. "Impact of Real-Time Satellite-Derived Green Vegetation Fraction on National Water Model Fluxes in Humid Alabama Watersheds." *Journal of the American Water Resources Association* 58(1): 104–18. https://doi.org/10.1111/1752-1688.12982.

Falcone, J.A. 2017. "U.S. Geological Survey GAGES-II Time Series Data from Consistent Sources of Land Use, Water Use, Agriculture, Timber Activities, Dam Removals, and Other Historical Anthropogenic Influences." U.S. Geological Survey Data Release. https://doi.org/10.5066/F7HQ3XS4.

Fall, G., D. Kitzmiller, S. Pavlovic, Z. Zhang, N. Patrick, M. St. Laurent, C. Trypaluk, W. Wu, and D. Miller. 2023. "The Office of Water Prediction's Analysis of Record for Calibration, Version 1.1: Dataset Description and Precipitation Evaluation." *Journal of the American Water Resources Association*: 1–27. https://doi.org/10.1111/1752-1688.13143.

Funakoshi, Y., J. Feyen, F. Aikman, H. Tolman, A. Van Der Westhuysen, A. Chawla, I. Rivin, and A. Taylor. 2012. "Development of Extratropical Surge and Tide Operational Forecast System (ESTOFS)." In *Proceedings of the International Conference on Estuarine and Coastal Modeling* 201–12.

Garnier, B.J., and A. Ohmura. 1968. "A Method of Calculating the Direct Shortwave Radiation Income of Slopes." *Journal of Applied Meteorology* 7: 796–800. https://doi.org/10.1175/1520-0450(1968)007<0796:AMOCTD>2.0.CO;2.

Garousi-Nejad, I., and D. Tarboton. 2022. "A Comparison of National Water Model Retrospective Analysis Snow Outputs at Snow Telemetry Sites across the Western United States." *Hydrological Processes* 2022(36): e14469. https://doi.org/10.1002/hyp.14469.

Givati, A., D. Gochis, T. Rummler, and H. Kunstmann. 2016. "Comparing One-Way and Two-Way Coupled Hydrometeorological Forecasting Systems for Flood Forecasting in the Mediterranean Region." *Hydrology.* 3(2): 19. https://doi.org/10.3390/hydrology3020019.

Gochis, D.J., M. Barlage, R. Cabell, M. Casali, A. Dugger, K. FitzGerald, M. McAllister, et al. 2020. "The WRF-Hydro® Modeling System Technical Description, (Version 5.1.1)." NCAR Technical Note, 107 pp. https://ral.ucar.edu/sites/default/files/public/WRFHydroV511TechnicalDescription.pdf.

Gochis, D.J., and F. Chen. 2003. "Hydrological Enhancements to the Community Noah Land Surface Model: Technical Description." NCAR Science and Technical Note, TN-454+STR, Sept., 2003.

Gourley, J.J., Z.L. Flamig, H. Vergara, P. Kirstetter, R.A. Clark, III, E. Argyle, A. Arthur, et al. 2017. "The FLASH Project: Improving the Tools for Flash Flood Monitoring and Prediction across the United States." *Bulletin of the American Meteorological Society* 98(2): 361–72.

Gower, J.C. 1971. "A General Coefficient of Similarity and Some of Its Properties." *Biometrics* 27(4): 857. https://doi.org/10.2307/2528823.

Hutton, E.W., M.D. Piper, and G.E. Tucker. 2020. "The Basic Model Interface 2.0: A Standard Interface for Coupling Numerical Models in the Geosciences." *Journal of Open Source Software* 5(51): 2317.

Julien, P.Y., B. Saghafian, and F.L. Ogden. 1995. "Raster-Based Hydrologic Modeling of Spatially-Varied Surface Runoff." *Journal of the American Water Resources Association* 31(3): 523–36. https://doi.org/10.1111/j.1752-1688.1995.tb04039.x.

Kitzmiller, D., D. Miller, R. Fulton, and F. Ding. 2013. "Radar and Multisensor Precipitation Estimation Techniques in National Weather Service Hydrologic Operations." *Journal of Hydrologic Engineering* 18(2): 133–42.

Lahmers, T.M., H. Gupta, C.L. Castro, D.J. Gochis, D. Yates, A. Dugger, D. Goodrich, and P. Hazenberg. 2019. "Enhancing the Structure of the WRF-Hydro Hydrologic Model for Semiarid Environments." *Journal of Hydrometeorology* 20: 691–714. https://doi.org/10.1175/JHM-D-18-0064.1.

Lahmers, T.M., P. Hazenberg, H. Gupta, C. Castro, D. Gochis, A. Dugger, D. Yates, L. Read, L. Karsten, and Y. Wang. 2022. "Evaluation of NOAA National Water Model Parameter Calibration in Semi-Arid Environments Prone to Channel Infiltration." *Journal of Hydrometeorology* 22(11): 2939–69.

Lehner, B., K. Verdin, and A. Jarvis. 2008. "New Global Hydrography Derived from Spaceborne Elevation Data." *Eos, Transactions, American Geophysical Union* 89(10): 93–94. https://doi.org/10.1029/2008eo100001.

Li, H., L. Luo, E.F. Wood, and J. Schaake. 2009. "The Role of Initial Conditions and Forcing Uncertainties in Seasonal Hydrologic Forecasting." *Journal of Geophysical Research* 114: D04114. https://doi.org/10.1029/2008JD010969.

Mason, L.A., A.D. Gronewold, M. Laitta, D.J. Gochis, K. Sampson, L. Read, E. Klyszejko, et al. 2019. "New Transboundary Hydrographic Data Set for Advancing Regional Hydrological Modeling and Water Resources Management." *Journal of Water Resources Planning and Management* 145: 06019004. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001073.

McKay, L., T. Bondelid, T. Dewald, J. Johnston, R. Moore, and A. Rea. 2012. *NHDPlus Version 2: User Guide*.

Nash, J.E., and J.V. Sutcliffe. 1970. "River Flow Forecasting through. Part I. A Conceptual Models Discussion of Principles." *Journal of Hydrology* 10: 282–90.

National Research Council. 2012. *Weather Services for the Nation: Becoming Second to None*. Washington, DC: National Academies Press.

National Weather Service (NWS). 2002. "Advanced Hydrologic Prediction Services—Concept of Services and Operations." Report, US Department of Commerce—NOAA—NWS, USA.

National Weather Service (NWS). 2022a. "NWS AHPS Forecast Website." United States Government. https://water.weather.gov/ahps/forecasts.php.

National Weather Service (NWS). 2022b. "NWS GFS Website." United States Government. https://www.emc.ncep.noaa.gov/emc/pages/numerical_forecast_systems/gfs.php.

National Weather Service (NWS). 2022c. "Weather Related Fatality and Injury Statistics." United States Government. https://www.weather.gov/hazstat/.

Niu, G.-Y., Z.-L. Yang, K.E. Mitchell, F. Chen, M.B. Ek, M. Barlage, A. Kumar, et al. 2011. "The Community Noah Land Surface Model with Multiparameterization Options (Noah-MP): 1. Model Description and Evaluation with Local-Scale Measurements." *Journal of Geophysical Research* 116(D12): 1–19. https://doi.org/10.1029/2010jd015139.

NOAA National Centers for Environmental Information (NCEI). 2022. "U.S. Billion-Dollar Weather and Climate Disasters." https://www.ncei.noaa.gov/access/billions/.

Nossent, J., and W. Bauwens. 2012. "Application of a Normalized Nash-Sutcliffe Efficiency to Improve the Accuracy of the Sobol's Sensitivity Analysis of a Hydrological Model." In: EGU General Assembly Conference Abstracts. p. 237.

OGC. 2017. "Open Geospatial Consortium, WaterML Version 2: Part 3—Surface Hydrology Features (HY_Features) Conceptual Model." Model standard. External identifier http://www.opengis.net/doc/IS/hy_features/1.0.

Ogden, F., B. Avant, R. Bartel, D. Blodgett, E. Clark, E. Coon, B. Cosgrove, et al. 2021. "The Next Generation Water Resources Modeling Framework: Open Source, Standards Based, Community Accessible, Model Interoperability for Large Scale Water Prediction." In AGU Fall Meeting Abstracts (Vol. 2021, pp. H43D-01). https://ui.adsabs.harvard.edu/abs/2021AGUFM.H43D..01O.

Pal, S., F. Dominguez, M.E. Dillon, J. Alvarez, C.M. Garcia, S.W. Nesbitt, and D. Gochis. 2020. "Hydrometeorological Observations and Modeling of an Extreme Rainfall Event Using WRF and WRF-Hydro during the RELAMPAGO Field Campaign in Argentina." *Journal of Hydrometeorology* 22: 331–51. https://doi.org/10.1175/JHM-D-20-0133.1.

Panofsky, H.A., and G.W. Brier. 1958. *Some Applications of Statistics to Meteorology*. University Park, PA: Earth and Mineral Sciences Continuing Education, College of Earth and Mineral Sciences, The Pennsylvania State University.

Paudel, Murari, E. James Nelson, Charles W. Downer, and Rollin Hotchkiss. 2011. "Comparing the Capability of Distributed and Lumped Hydrologic Models for Analyzing the Effects of Land Use Change." *Journal of Hydroinformatics* 13(3): 461–73.

Read, L.K., D.N. Yates, J.M. McCreight, A. Rafieeinasab, K. Sampson, and D.J. Gochis. 2023. "Development and Evaluation of the Channel Routing Model and Parameters within the National Water Model." *Journal of the American Water Resources Association* 59(5): 1051–66. https://doi.org/10.1111/1752-1688.13134.

Reed, S., V. Koren, M. Smith, Z. Zhang, F. Moreda, D.J. Seo, and DMIP Participants. 2004. "Overall Distributed Model Intercomparison Project Results." *Journal of Hydrology* 298(1–4): 27–60. https://doi.org/10.1016/j.jhydrol.2004.03.031.

Restrepo, P., J. Roe, C. Dietz, M. Werner, P. Gijsbers, R. Hartman, H. Opitz, B. Olsen, J. Halquist, and R. Shedd. 2010. "Hydrologic Forecasting at the US National Weather Service in the 21st Century: Transition from the NWS River Forecast System (NWSRFS) to the Community Hydrologic Prediction System (CHPS)." EGU General Assembly Conference Abstracts, p. 6760.

Rogers, E., J.R. Carley, B.S. Ferrier, E. Aligo, G. Gayno, Z. Janjic, Y. Lin, et al. 2017. "Mesoscale Modeling Development at the National Centers for Environmental Prediction: Version 4 of the NAM Forecast System and Scenarios for the Evolution to a High-Resolution Ensemble Forecast System." 28th Conf. on Weather Analysis and Forecasting/24th Conf. on Numerical Weather Prediction, Seattle, WA, Amer. Meteor. Soc., 3B.4. https://ams.confex.com/ams/97Annual/webprogram/Paper311212.html.

Rummler, T., J. Arnault, D.J. Gochis, and H. Kunstmann. 2019. "Role of Lateral Terrestrial Water Flow on the Regional Water Cycle in a Complex Terrain Region: Investigation with a Fully Coupled Model System." *Journal of Geophysical Research: Atmospheres* 124: 507–29. https://doi.org/10.1029/2018JD029004.

Saha, S., S. Moorthi, X. Wu, J. Wang, S. Nadiga, P. Tripp, D. Behringer, et al. 2014. "The NCEP Climate Forecast System Version 2." *Journal of Climate* 27(6): 2185–208.

Senatore, A., L. Furnari, and G. Mendicino. 2020. "2020: Impact of High-Resolution Sea Surface Temperature Representation on the Forecast of Small Mediterranean catchments' Hydrological Responses to Heavy Precipitation." *Hydrology and Earth System Sciences* 24: 269–91. https://doi.org/10.5194/hess-24-269-2020.

Senatore, A., G. Mendicino, D.J. Gochis, W. Yu, D.N. Yates, and H. Kunstmann. 2015. "Fully Coupled Atmosphere-Hydrology Simulations for the Central Mediterranean: Impact of Enhanced Hydrological Parameterization for Short and Long Time Scales." *Journal of Advances in Modeling Earth Systems* 7: 1693–715. https://doi.org/10.1002/2015MS000510.

Skamarock, W.C., J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, M.G. Duda, X.Y. Huang, W. Wang, and J.G. Powers. 2008. "A Description of the Advanced Research WRF Version 3." NCAR Tech. Note NCAR/TN-4751STR. http://www2.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf.

Sweeney, T. 1992. "Modernized Areal Flash Flood Guidance." NOAA Tech. Rep. NWS HYDRO 44, NOAA/NWS/Hydrologic Research Laboratory, Silver Spring, MD, 21 pp. + an Appendix.

Taylor, A., and B. Glahn. 2008. "Probabilistic Guidance for Hurricane Storm Surge." Preprints, 19th Conference on Probability and Statistics, 88th Amer. Meteor. Soc. Annual Meeting, New Orleans, LA, 7.4., 8 pp.

Tijerina, D., L. Condon, K. FitzGerald, A. Dugger, M. Forrester, K. Sampson, D. Gochis, and R. Maxwell. 2021. "Continental Hydrologic Intercomparison Project (CHIP), Phase 1: A Framework for Large-Scale Hydrologic Model Comparisons." *Water Resources Research* 57. https://doi.org/10.1029/2020WR028931.

Tolson, B.A., and C.A. Shoemaker. 2007. "Dynamically Dimensioned Search Algorithm for Computationally Efficient Watershed Model Calibration." *Water Resour. Res.* 43: W01413. https://doi.org/10.1029/2005WR004723.

Valayamkunnath, P., D.J. Gochis, F. Chen, M. Barlage, and K.J. Franz. 2022. "Modeling the Hydrologic Influence of Subsurface Tile Drainage Using the National Water Model." *Water Resources Research* 58(4): 1–24. https://doi.org/10.1029/2021wr031242.

Verri, G., N. Pinardi, D. Gochis, J. Tribbia, A. Navarra, G. Coppini, and T. Vukicevic. 2017. "A Meteo-Hydrological Modelling System for the Reconstruction of River Runoff: The Case of the Ofanto River Catchment." *Natural Hazards and Earth System Sciences* 17: 1741–61. https://doi.org/10.5194/nhess-17-1741-2017.

Vionnet, Vincent, E. Brun, S. Morin, S. Faroux, P. Le Moigne, E. Martin, and J. Willemet. 2012. "The Detailed Snowpack Scheme Crocus and its Implementation in SURFEX v7.2." *Geoscientific Model Development*. 5: 773–91. https://doi.org/10.5194/gmd-5-773-2012.

Viterbo, F., K. Mahoney, L. Read, F. Salas, B. Bates, J. Elliott, B. Cosgrove, D.J. Gochis, and R. Cifelli. 2020. "A Multiscale, Hydrometeorological Forecast Evaluation of National Water Model Forecasts of the May 2018 Ellicott City, Maryland, Flood." *Journal of Hydrometeorology* 21: 475–99. https://doi.org/10.1175/JHM-D-19-0125.1.

Viterbo, F., L. Read, K. Nowak, A.W. Wood, D.J. Gochis, R. Cifelli, and M. Hughes. 2020. "General Assessment of the Operational Utility of National Water Model Reservoir Inflows for the Bureau of Reclamation Facilities." *Water* 12(10). https://doi.org/10.3390/w12102897.

Wigmosta, M., L. Vail, and D. Lettenmaier. 1994. "A Distributed Hydrology-Vegetation Model for Complex Terrain." *Water Resources Research* 30: 1665–79.

Wood, A.W., E.P. Maurer, A. Kumar, and D.P. Lettenmaier. 2002. "Long-Range Experimental Hydrologic Forecasting for the Eastern United States." *Journal of Geophysical Research* 107(D20): 4429. https://doi.org/10.1029/2001JD000659.

Xia, Y., K. Mitchell, M. Ek, J. Sheffield, B. Cosgrove, E.F. Wood, L. Luo, et al. 2012. "Continental-Scale Water and Energy Flux Analysis and Validation for the North American Land Data Assimilation System Project Phase 2 (NLDAS-2): 1. Intercomparison and Application of Model Products." *Journal of Geophysical Research Atmospheres* 117(3): D03109. https://doi.org/10.1029/2011JD016048.

Xiang, T., E.R. Vivoni, D.J. Gochis, and G. Mascaro. 2017. "On the Diurnal Cycle of Surface Energy Fluxes in the North American Monsoon Region Using the WRF-Hydro Modeling System." *Journal of Geophysical Research—Atmospheres*. 122(17): 9024–49.

Yang, Z.-L., G.-Y. Niu, K.E. Mitchell, F. Chen, M.B. Ek, M. Barlage, L. Longuevergne, et al. 2011. "The community Noah land Surface Model with Multiparameterization Options (Noah-MP): 2. Evaluation over Global River Basins." *Journal of Geophysical Research* 116(D12): 1–16. https://doi.org/10.1029/2010jd015140.

Yucel, I., A. Onen, K. Yilmaz, and D. Gochis. 2015. "Calibration and Evaluation of a Flood Forecasting System: Utility of Numerical Weather Prediction Model, Data Assimilation and Satellite-Based Rainfall." *Journal of Hydrology* 523: 49–66.

Zangl, G. 2005. "Formation of Extreme Cold-Air Pools in Elevated Sinkholes: An Idealized Numerical Process Study." *Monthly Weather Review* 133: 925–41. https://doi.org/10.1175/MWR2895.1.

Zhang, J., K. Howard, C. Langston, B. Kaney, Y. Qi, L. Tang, H. Grams, et al. 2016. "Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation: Initial Operating Capabilities." *Bulletin of the American Meteorological Society* 97(4): 621–38.

Zhang, Y., F. Ye, E.V. Stanev, and S. Grashorn. 2016. "Seamless Cross-Scale Modeling with SCHISM." *Ocean Modelling* 102: 64–81.

Zhou, Q., G. Leng, and J. Peng. 2018. "Recent Changes in the Occurrences and Damages of Floods and Droughts in the United States." *Water* 10(9): 1109.

How to cite this article: Cosgrove, Brian, David Gochis, Trey Flowers, Aubrey Dugger, Fred Ogden, Tom Graziano, Ed Clark et al; 2024. "NOAA's National Water Model: Advancing Operational Hydrology Through Continental-scale Modeling." *JAWRA Journal of the American Water Resources Association* 60 (2): 247–272. https://doi.org/10.1111/1752-1688.13184.

APPENDIX A

A.1 | National Water Model analysis configurations

The main baseline analysis and assimilation (AnA) cycle initializes the short- and medium-range forecasts cycles hourly and has a three-hour lookback period. This lookback period begins 3 h before the current cycle time, and progresses up to the current cycle time (i.e., the 12Z AnA begins at 09Z and ends at 12Z). The benefit of the lookback period is that it enables the model to ingest observation- in addition to forecast-based precipitation forcing data. This forcing data are drawn preferentially from the Multi-Radar Multi-Sensor (MRMS) gauge-adjusted product, with the MRMS radar-only observed precipitation product used if the primary source is unavailable. Short-range Rapid Refresh (RAP) and High-Resolution Rapid Refresh (HRRR) forecasts are used in areas or for times where MRMS radar coverage is poor. One-hour RAP and HRRR forecasts supply the other meteorological forcing information. Real-time USGS streamflow observations are assimilated into this configuration via the nudging approach discussed previously. Taken together, this system produces a real-time snapshot of the current streamflow and general hydrologic states across the country which can be used both to initialize National Water Model (NWM) short- and medium-range forecasts and as input to various end user applications.

Once per day, this baseline AnA is supplemented by an extended lookback AnA. The 28-h lookback period of this alternate configuration allows for the ingestion of higher-quality precipitation data from the National Centers for Environmental Prediction Stage IV precipitation dataset. This dataset is a CONUS mosaic of the separate observation-based multisensor precipitation estimator (MPE) grids produced by the 12 CONUS River Forecast Centers (RFCs) (Kitzmiller et al., 2013). The use of this high-quality product by the Extended AnA increases the accuracy of the resulting hydrologic simulations, and also promotes operational consistency with the RFCs which use MPE precipitation data in their operational river forecasting activities. With its 28-h lookback period, this simulation runs from 12Z the previous day, to 16Z the current day. Model states from the end of this simulation replace the baseline AnA states valid at that same time, ensuring that once per day there is an injection of higher quality states from this independently cycling simulation.

The third CONUS AnA configuration is the Long-Range AnA used to initialize NWM Long-Range forecasts. Executed four times per day with a 12 h lookback period, this configuration differs further from the other two CONUS AnA cycles via its use of the simplified approach to routing surface water discussed earlier. Configured in this way to reduce execution time, this approach mirrors the physics configuration within the NWM Long-Range forecast. The same set of forcing data is used in the Long-Range AnA as in the baseline AnA, with the same set of USGS observations assimilated as well.