



Editor-in-Chief
Giles Hollister
Associate Editors
Steven Quinlan, Christine Lortie, Theodore L. Jordan-Morgan

An analysis of the National Water Model for a mid-Atlantic suburban watershed

Saber E. Brasher, Daniel J. Leathers, Christina L. Callahan & Kathryn E. Giesa

To cite this article: Saber E. Brasher, Daniel J. Leathers, Christina L. Callahan & Kathryn E. Giesa (2024) An analysis of the National Water Model for a mid-Atlantic suburban watershed, *Physical Geography*, 45:1, 84-105, DOI: [10.1080/02723646.2023.2260546](https://doi.org/10.1080/02723646.2023.2260546)

To link to this article: <https://doi.org/10.1080/02723646.2023.2260546>



Published online: 21 Sep 2023.



Submit your article to this journal [↗](#)



Article views: 222



View related articles [↗](#)



View Crossmark data [↗](#)



An analysis of the National Water Model for a mid-Atlantic suburban watershed

Saber E. Brasher ^a, Daniel J. Leathers ^b, Christina L. Callahan^b and Kathryn E. Giesa^b

^aDepartment of Geography, Oklahoma State University, Stillwater, OK, USA; ^bDepartment of Geography and Spatial Sciences, University of Delaware, Newark, DE, USA

ABSTRACT

Discharge values from the National Water Model (NWM) were compared to USGS stream gage discharge observations for the suburban Red Clay Creek watershed (drainage area ~140 km² and mixed land-use), in Pennsylvania and Delaware, from 2016 to 2018. 18-hour retrospective simulations from the NWM were used with concurrent hourly USGS discharge observations from three locations along the Red Clay Creek. Results indicate that the mean of discharge estimates from the NWM and from USGS observations significantly differed and that the NWM generally underestimates low-flow conditions and overestimates high-flow conditions. Watershed size also impacted NWM performance (with performance degrading in smaller watersheds). A meteorological analysis determined that convective rainfall events were associated with 66% of the largest differences between NWM discharge estimates and USGS observations while mid-latitude cyclone stratiform precipitation events accounted for the other 34%. Lastly, of the largest 15 differences between the NWM and observations, 13 occurred with pre-cursor soil moisture that was below the mean (dry soil conditions), in conjunction with heavy rainfall. Given the NWM's recent operational implementation, and its status as Prototype guidance, the results of this study present specific geographical and climatological findings that can aid in the NWM's continued validation and improvement for similar regions.

ARTICLE HISTORY

Received 21 March 2023

Accepted 26 August 2023

KEYWORDS

National water model; streamflow; mid-atlantic; watershed analysis; Climate; water Resources

Introduction

Accurate predictions of streamflow are essential in regions where flooding is common and where surface streams are a major contributor to water resources. This is the case in much of the mid-Atlantic region of the United States where large population centers depend on surface water for consumption and sanitation, and where both topographically induced and urban flooding often take place (Smith & Smith, 2015; Dieter et al., 2018). The Christina River Basin in southeast Pennsylvania and northern Delaware comprises four watersheds (White Clay Creek, Red Clay Creek, Brandywine Creek, and Christina River) across its 1463 km² area. The basin has been subject to major

flooding in the past and contributes more than 50% of the water supply for a population of more than 500,000 people (Cruz-Ortiz and Miller, 2013).

A tributary to the Christina River, Red Clay Creek, drains approximately 10% of the Christina basin. The Red Clay Creek watershed was severely inundated in 2003 as the remnants of Tropical Storm Henri flooded the area, devastating the community of Glenville, DE. During this event, more than 10 inches of rain fell within a five-hour period, exceeding the 100-year return value for a 24-hour event, with discharge values beyond $900 \text{ m}^3/\text{s}$ (in a small drainage area of only about 140 km^2 ; source: weather.gov). In addition, water from Red Clay Creek is used for human consumption and supports two public water supply systems (Cruz-Ortiz and Miller 2013). To better prepare for extreme flooding events and ensure adequate local water availability a thorough understanding of the current tools for predicting streamflow on the Red Clay Creek is needed. To meet this objective, this manuscript will compare the relative performance of the National Water Model (NWM), the standard for streamflow forecasting in the United States, to U.S Geological Survey (USGS) streamflow observations in three separate reaches within the Red Clay Creek from 2016 to 2018.

Recent warming trends across the mid-Atlantic region will likely impact both water resources and the flood hydroclimatology of the region. As the planet warms, there is a direct impact on precipitation as the warmer temperatures allow increased evaporation and an increase in the saturation vapor pressure of the atmosphere by about 7% for each 1°C of global warming (Griffiths & Bradley, 2007; Trenberth, 2011). This increase in precipitable water allows for storms with more intense precipitation rates. The mid-Atlantic has already experienced precipitation totals exceeding the 99th percentile for daily amounts increasing significantly from 1957 to 2010 (Kunkel et al., 2013). Kunkel et al. (2013) also documented an increasing amount of rainfall in the Northeast specifically attributed to tropical cyclones, not because of a significant increase in the number of storms, but because of higher precipitation totals associated with each event. Across the United States, most trends in the 2-, 5-, and 10-year return period rainfall amounts are positive, implying an increased number of intense storms. This trend was also found to be statistically significant for the Northeast United States (Degaetano, 2009). Moreover, Leathers et al. (2020) used Delaware's high-resolution weather-network to investigate the return periods of storm events and found an increase in the number of long-duration, high intensity precipitation events, most of which were associated with tropical or mid-latitude cyclones (Leathers et al., 2020). The area studied in Leathers et al. (2020) includes the Red Clay Creek Basin. With increased variance in hydroclimatic events, accurately measuring the impact on streamflow will be imperative for applications in both disaster forecasting (flooding) as well as water resource availability for consumption.

The National Oceanic and Atmospheric Administration's (NOAA) National Water Model (NWM) became operational in 2016 and is the United States' most extensive water prediction tool, delivering forecasts for about 2.7 million stream reaches over the entire continental United States. Forecasts from the NWM are considered prototype guidance by the National Weather Service (NWS) Office of Water Prediction (OWP) and are not an "official" NWS river forecast. However, as the model matures and is made more readily available to decision makers and the public, its use as a forecast tool is likely to expand. The model is based on the National Center for Atmospheric Research's (NCAR) WRF-Hydro model (Bales & Flowers, 2021). WRF-Hydro ingests data from a variety of

sources including multiple NOAA weather prediction models (HRRR, GFS, NAM-Nest, etc.) and NOAA's Multi-Radar/Multi-Sensor System (MRMS) and is essentially a large physically (or process) driven model. Output of the NWM includes: 1. Short-range forecasts that are made 18 hours into the future and are deterministic including a single forecast value, 2. Medium-range forecasts comprising seven-member ensembles ranging from 8.5 to 10 days into the future, and 3. Long-range forecasts that run out to 30 days as a four-member ensemble. Given its recent operational implementation, and its status as Prototype guidance, the model continues to be validated for accuracy in different regions and under varying scenarios for continued improvement.

Although many studies focus on flood forecasting, it has recently been discovered that the NWM struggles to perform well in low-flow conditions, as the model only represents streamflow exchange with groundwater as a one-way exchange (moving from an aquifer into the stream, but not the stream into the aquifer) (Bales & Flowers, 2021) and may overestimate streamflow in regions with losing reaches (those that have stream water return to aquifers) while underestimating flood magnitudes in gaining reaches (those that gain water directly from aquifers) (Jachens et al., 2021). It was found that the NWM underestimated low-flows, locations with low-flows, and the duration of low-flow events in the Colorado River Basin and Northern High Plains region especially during severe or exceptional drought (Hansen et al., 2019; Karki et al., 2021). The underestimation of low-flow has also been linked to water velocity inaccuracy, and the NWM performs better in forested landcover types than urbanized ones (Duan & Kumar, 2020; Duan et al., 2023). Lastly, the NWM has been documented as performing best in large watershed basins and less well in smaller basins (Rojas et al., 2020) and observation accuracy continuity throughout the stream (where high-quality observations upstream influence model output downstream) also improves results (Rojas et al., 2020).

An evaluation of the ability of the NWM to predict the Ellicott City, MD flooding event of May 2018 was conducted by Viterbo et al. (2020) where the authors performed a multiscale hydrometeorological forecast assessment of the overall efficacy of a nationally distributed modeling approach such as the NWM. The study suggested that there were potential advantages in the use of the NWM but also risks due to the complexity of the inputs to the model. The utility of NWM forecasts for estimating reservoir inflows was evaluated by Viterbo et al. (2020) where the authors found that the NWM offers inflow information for over 5000 reservoirs across the United States. Their results indicate that the NWM performs better in snow-driven inflow basins than in rain-driven basins. In addition, the basin area and stream management practices are both important in the accuracy of the daily reservoir inflows. Finally, Duan and Kumar (2020) investigated the usefulness of the NWM in seasonal streamflow predictability in the Alabama-Coosa-Tallapoosa (ACT) River Basin in the southeastern United States. The authors found that the NWM accurately predicted streamflow in the basin, and that initial soil moisture condition is a major factor in the predictability of the system on seasonal timescales.

To summarize, the NWM, although an innovative product that has many applications and performs well under many scenarios, requires more validation and refinement under specific hydrologic settings and hydroclimatic conditions to be trusted as a reliable resource. It is the purpose of this research to evaluate the performance of the NWM in a mid-Atlantic suburban watershed, with an emphasis on the meteorological conditions during high-flow events in the context of a small, mixed land use basin.

Study area

The Red Clay Creek watershed, located within the Mid-Atlantic region of the United States (Figure 1), straddles the border of Pennsylvania and Delaware draining an area of approximately 140 km² (Senior & Koerkle, 2003). The headwaters of the stream are in Chester County, PA with the creek flowing generally south into New Castle County, DE. The basin is home to approximately 48,000 people (Cruz-Ortiz and Miller, 2013) and it supplies drinking water for residents of Delaware.

Elevations in the basin range from approximately 160 m in the northern portion to approximately 1 m at the confluence with White Clay Creek at Stanton, DE (Figure 2b). Forest and other vegetated surfaces make up the majority of land cover across the basin (Figure 2b), while developed surfaces comprise the next greatest land coverage (Table 1). The basin includes the Piedmont physiographic province in southeastern Pennsylvania and the Piedmont and Coastal Plain province in northern Delaware, where the topography can be categorized as transitioning between gently rolling hills with narrow valleys to nearly flat terrain (Senior & Koerkle, 2003). Senior and Koerkle (2003) describe the rock type beneath the basin as primarily metamorphic and sedimentary in origin, with about 90% of the soil association being Glenelg-Manor-Chester, which is typically well drained, ranging in permeability from about 0.6 to 2.0 in/hr. in most cases. Streamflow readily responds to the amount of impervious cover nearby, with studies documenting higher runoff peaks and volume as landscape development increases (Shuster et al., 2005). It was documented that the “flashiness” (or the sum of the absolute values of day-to-day changes in mean daily flow divided by the sum of the mean daily flow) was positively correlated

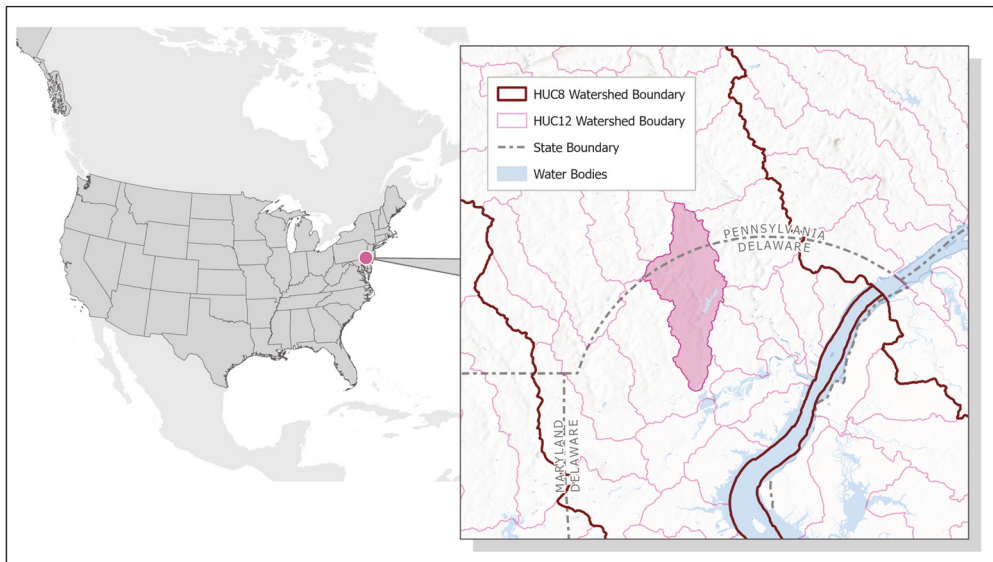


Figure 1. Map showing the location of the Red Clay Creek watershed (highlighted in magenta), straddling the border between Delaware and Pennsylvania, in the mid-Atlantic region of the United States.

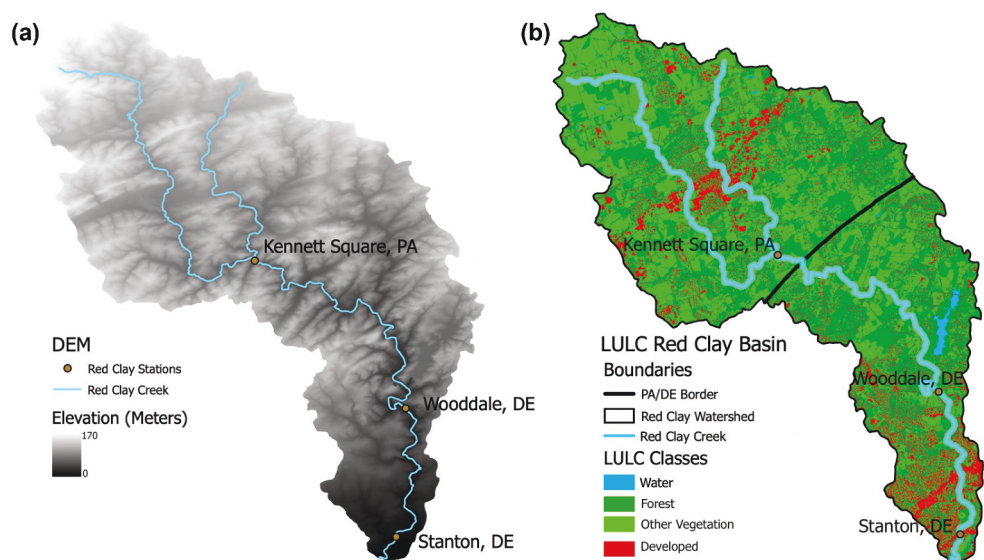


Figure 2. Maps showing a) elevation and the location of USGS stream gaging stations in the Red Clay Creek basin (data: <https://www.Dgs.udel.edu/datasets/elevation-contours-delaware>), and b) simplified land use and land cover classification (where “developed” is impervious covers and “other vegetation” includes other non-developed covers and agriculture) in the basin (data: <https://de-firstmap-delaware.hub.arcgis.com/pages/data>).

Table 1. Distribution of LULC class coverage across the three study basins (where “developed” is developed lands and impervious covers and “other vegetation” includes remaining basin land covers), with percentages adapted from Hammond et al. (2022)..

Location	Water (%)	Forest (%)	Developed (%)	Other Vegetation (%)
Kennett Square, PA	2.40	26.74	4.98	29.88
Wooddale, DE	3.26	36.5	34.46	25.78
Stanton, DE	3.47	34.50	4.26	21.77

with increased impervious cover and was often highest in smaller stream areas (Moltz et al., 2018). In the nearby (directly west of the Red Clay Creek watershed) White Clay Creek watershed, correlations suggest a decrease in the stream’s base flow with an increase in impervious cover, attributing the decrease to a loss in permeable recharge areas (with vegetated landscapes replaced by more pavement and structures), as well as water being intercepted by storm drains and sewers (Kauffman et al., 2009).

These watershed-specific details could greatly contribute to streamflow modeling accuracy and impact decision support utilization of the NWM. The Red Clay basin is one of many in the mid-Atlantic region that contains a blend of mixed forest, agricultural and urban or suburban land use; within the Piedmont physiographic province, ~33% of watershed land use classes are either urban or mixed cover with some vegetation (Wardrop et al., 2005). As such, we expect that the results of this comparison of the NWM and observed streamflow are representative of other similar basins across the mid-Atlantic region, thus making the Red Clay basin a valuable watershed for analysis.

Data and methods

To test the NWM’s performance in the Red Clay Creek basin, this study compares United States Geological Survey (USGS) observation data from three reaches on the Red Clay Creek to comparable 18-hour NWM retrospective simulation data. USGS stream discharge data is extensively quality controlled and will be used as the “control” for this study. However, it is important to note that USGS discharge measurements are based on ratings curves that depend upon the physical characteristics of the stream channel and floodplain, both of which are changing over time (USGS, 2010). Thus, stream discharge values are “estimates” of the flow based upon the accuracy of the ratings curve for a particular time and location (Turnipseed & Sauer, 2010).

The USGS stream gage stations and corresponding NWM reach IDs (Table 2) used are from Kennett Square, Pennsylvania (the northernmost station used), Wooddale, Delaware (downstream of the Kennett Square station), and Stanton, Delaware (the southernmost station used) (Figure 2a). The drainage area increases moving south from Kennett Square, PA (73 km²) to Wooddale, DE (122 km²) and finally Stanton, DE (136 km²). All locations have a water stage recorder and crest-stage gage and use the stage–discharge relation to convert the continuously measured stream stage measurements into an estimate of streamflow (Turnipseed & Sauer, 2010). Hourly discharge observations for a three-year period (2016–2018) were obtained for the three USGS stations described above from the USGS Water Data for the Nation NWIS site (U.S Geological Survey, 2023; <https://waterdata.usgs.gov/nwis>). Version 2.0 NWM 18-hour retrospective simulations for the reaches corresponding to the USGS gaging stations were obtained from the Office of Water Prediction AWS site (NOAA, 2023; <https://registry.opendata.aws/nwm-archive/>) and used for this analysis. Version 2.0 of the NWM did not assimilate stream gage observations.

Precipitation and soil volumetric water content (VWC) data were obtained from the Delaware Environmental Observing System (DEOS). DEOS precipitation measurements are made using a Texas Electronics TE-525USW tipping bucket rain gauge, while volumetric water content data is collected using a Campbell Scientific CS655 or CS616 water content reflectometer. Data from two stations in the watershed, Kennett Square (henceforth “Bucktoe” as there are two DEOS stations in Kennett Square), PA and Mt. Cuba, DE were used in this study.

Multiple methods were used to compare the USGS stream discharge data and the NWM estimates for the corresponding stream reach. Due to the positive skewness of discharge values, both the mean and median were considered as measures of central tendency throughout the analysis. A student’s t-test and Z-test were used to determine if the means of the observed and modeled discharge were from similar distributions. Although the discharge values are not strictly normally distributed, for the large sample

Table 2. USGS IDs and latitude/longitude coordinates for the three testing locations used in the study along with the corresponding NWM reach IDs.

Location	NWM ID	USGS ID	USGS Coordinates
Kennett Square, PA	4651090	01479820	39.81677739, –75.6916008
Wooddale, DE	4651912	01480000	39.76280556, –75.6365
Stanton, DE	4651930	01480015	39.71575, –75.6399444

sizes used in this study the t-test is a robust statistic (Davis, 1986). To further verify the results of the t and Z tests, the non-parametric Wilcoxon Signed Rank test was also used in the case of non-normal distributions (Davis, 1986). To examine the agreement of the observed and modeled discharge pairs, simple least squares linear regression was used. In general, R^2 values are used in this analysis to understand the proportion of variance explained in the discharge observations by the modeled discharge. Cumulative probability curves were calculated to determine the frequency of occurrence of values of discharge less than a given reference value for both observed and modeled discharge. These curves are used to better understand the tendency of the model to under-or-over-predict flow values at each station. Finally, basic climate compositing analysis, calculating the spatial mean of a meteorological variable for a given number of events, was used to produce the maps in the meteorological analysis.

Results

Results include a discussion of the discharge characteristics at each USGS gaging station in addition to how the model values compared to the observation values. An analysis of the meteorological conditions during extreme flooding events, including the categorization of the type of precipitation that took place, is also detailed below.

Discharge characteristics at USGS gaging stations

A comparison of stream discharge for each of the three USGS gaging stations located along Red Clay Creek in Pennsylvania and Delaware for the period of record (2016–2018) was conducted. The annual median discharge at each location is shown in Figure 3a. As expected from the size of their watersheds, the largest median discharge is found at Stanton, DE each year, followed by Wooddale, DE and Kennett Square, PA. However, the relative difference between the median value changes from one year to another, and median values for all locations are larger in 2018 compared to the previous two years.

The annual cycle of median discharge is similar at all stations (Figure 3b) with the largest values occurring during the spring months, and the lowest discharge during the autumn season. A time series of discharge for all stations for an active discharge period (November 1–15, 2018) is shown in Figure 4. The three stations respond similarly to precipitation inputs in proportion to their basin areas. Lag times in peak flows between stations are generally short, approximately two hours from Kennett Square to Wooddale and less between Wooddale and Stanton.

Comparison of model with observations

Summary statistics for each gaging station are given in Table 3 for both the observed and modeled discharge. For all stations, modeled means and standard deviations were both larger than observed values, while modeled median values were smaller than the observed. This suggests that the model is overestimating large discharge values and underestimating small discharge values.

To test for differences in the distribution of the modeled versus observed data for each location a t-test and Z-test were used, in addition to a non-parametric

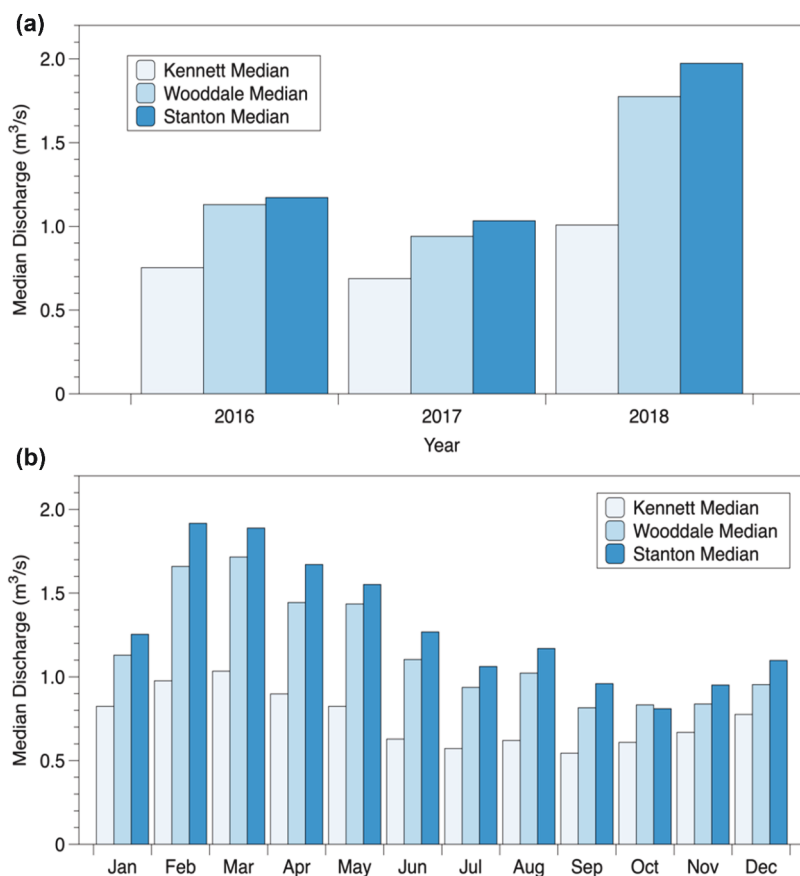


Figure 3. Annual median discharge (a) and monthly median discharge (b) for the three USGS gaging stations (Kennett ID: 01479820, Wooddale ID: 01480000, Stanton ID: 01480015) along Red Clay Creek from 2016–2018. Discharge given in m^3/s .

Wilcoxon Signed Rank test. The null hypothesis was rejected for each test for each location ($p < 0.001$) suggesting that the modeled and observed data are from different distributions. The data were further separated by discharge value to include only the top 10% of values at each station (those that would most likely be associated with flooding and societal disruption). Again, the null hypothesis was rejected for each location ($p < 0.001$) indicating that the modeled and observed data were likely from two separate distributions for the top 10% of discharge values.

To further examine the correspondence of the observed and modeled discharge, simple least squares linear regression was used. Figure 5 shows scatter plots of the relationship between observed and modeled discharge at each station, along with the least squares line, and the coefficient of variation (R^2). The value of R^2 is smallest at Kennett Square, PA ($R^2 = 0.37$, 5a; the smallest basin) and increases as the basin increases in size moving south to Wooddale, DE ($R^2 = 0.51$, 5b) and finally Stanton, DE ($R^2 = 0.54$, 5c).

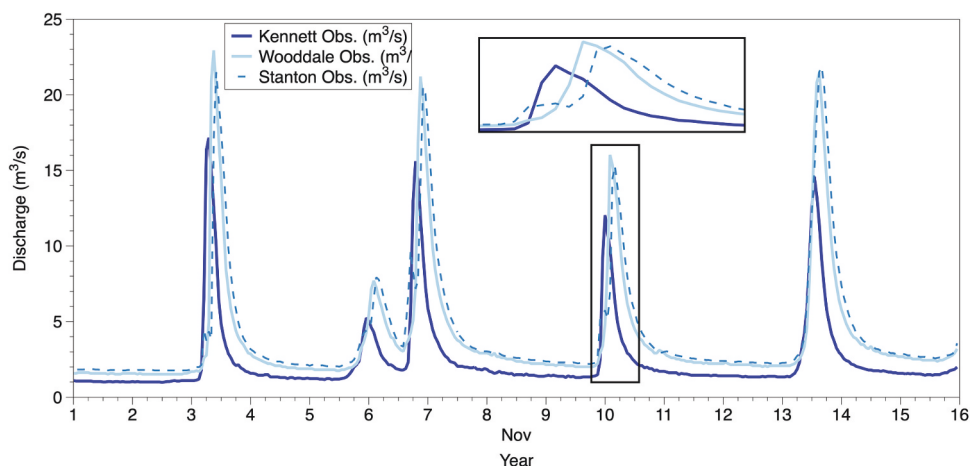


Figure 4. Time series of hourly discharge values for the three USGS gaging stations along Red Clay Creek. Kennett Square, PA (USGS ID: 01479820), Wooddale, DE (USGS ID: 01480000), and Stanton, DE (USGS ID: 01480015), for November 1–15, 2018. Values given in units of (m^3/s). Inset shows lags in peak flow for a single event.

Table 3. Summary statistics (mean, standard Deviation, median, and number of observations) for observed (USGS) and modeled (NWM) values for each gaging station for three years (2016–2018) of hourly values.

Statistic	Kennett Square		Wooddale		Stanton	
	Observed	Modeled	Observed	Modeled	Observed	Modeled
Mean	1.16 m^3/s	1.41 m^3/s	1.85 m^3/s	2.25 m^3/s	2.05 m^3/s	2.52 m^3/s
S.D.	2.21 m^3/s	3.56 m^3/s	3.60 m^3/s	5.32 m^3/s	3.93 m^3/s	5.47 m^3/s
Median	0.80 m^3/s	0.66 m^3/s	1.22 m^3/s	1.00 m^3/s	1.35 m^3/s	1.21 m^3/s
N	26109	25862	25333	25526	25204	25862

A graph showing the mean monthly NWM and observed discharge for Wooddale, DE (middle of the basin and middle-sized watershed area) is shown in [Figure 6a](#). Each month, except for April, shows higher mean discharge in the NWM compared to observations. Differences are more notable in some months (February, March, July, November), and all gaging stations show similar annual cycle differences. [Figure 6b](#) shows the annual cycle of median observed and modeled discharge for Wooddale, DE. In this case, the modeled median discharge values are smaller than observed for all months except February and March, suggesting a possible connection to snow cover ablation or frozen soils within the model increasing the median discharge during those two months.

Differences in the distributions between modeled and observed data are clear with inspection of cumulative probability curves for each station ([Figure 7a-c](#)). At each location the NWM generally underestimates discharge compared to observations for low-flow events ($<1 \text{ m}^3/\text{s}$) and overestimates discharge for events greater than $1 \text{ m}^3/\text{s}$.

The bias of the model for overestimating large discharge values is illustrated in [Figure 8](#), which is a time series of the modeled versus observed data for the period of

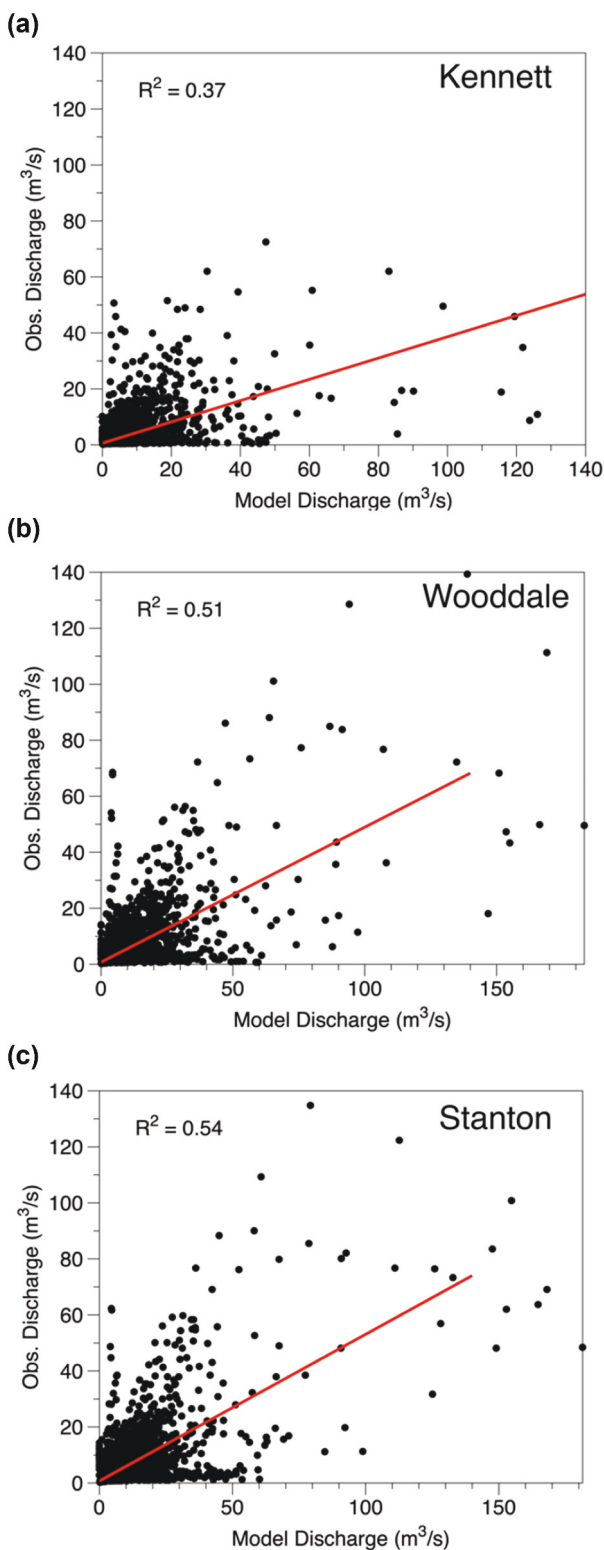


Figure 5. Scatter plots showing the relationship between observed (USGS) and modeled (NWM) discharge at Kennett Square, PA (a), Wooddale, DE (b), and Stanton, DE (c). The red line is the least squares line, and the coefficient of variation is given for each scatterplot.

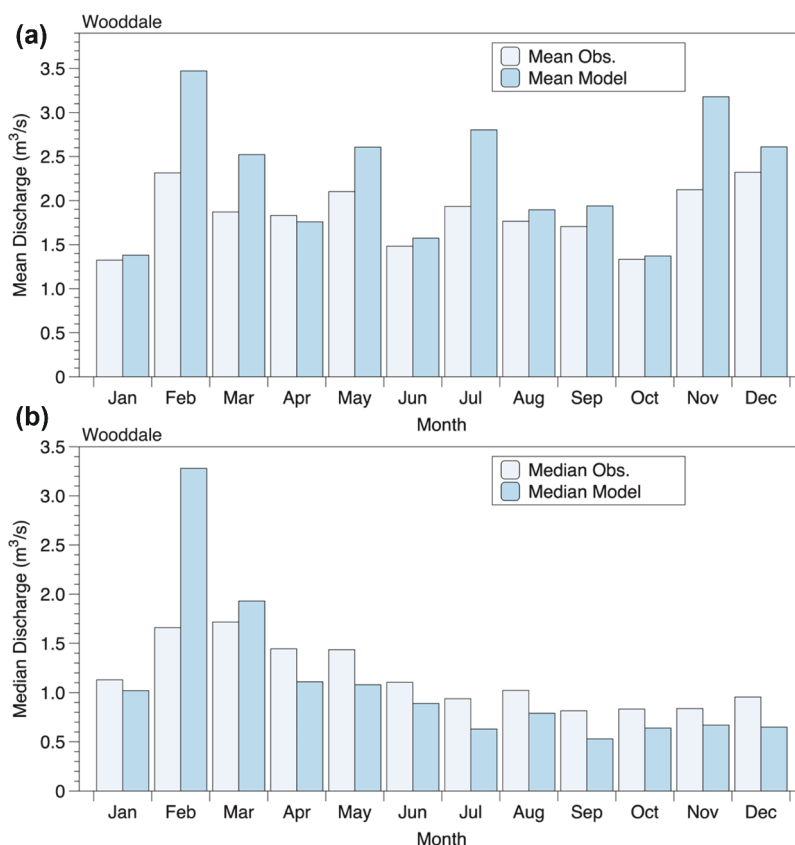


Figure 6. Annual cycle of a) mean and b) median observed (USGS) and modeled (NWM) discharge for Wooddale, DE. Values are given in (m^3/s).

record at Wooddale, DE (mid-basin). In nearly all large discharge events, the model is significantly overestimating the flow at Wooddale. These overestimates by the model would, at times, be suggesting flooding conditions along the Red Clay Creek when none occurred or would forecast more severe flooding conditions than took place (the blue line represents the onset of flood conditions at Wooddale, DE).

Analysis of extreme events

The performance of the NWM during flooding events is of particular importance for societal concerns. Over-prediction or under-prediction by the model, at high observed flows, presents an issue in short-range flood forecasting. Analysis of the model's response under diverse atmospheric conditions could assist users in evaluating the limitations and tendencies of its performance. Moreover, examining the meteorological situation during instances when the model performs poorly (largest differences between the modeled and observed discharge) may give additional guidance on the most appropriate use cases for the NWM in a small suburban basin.

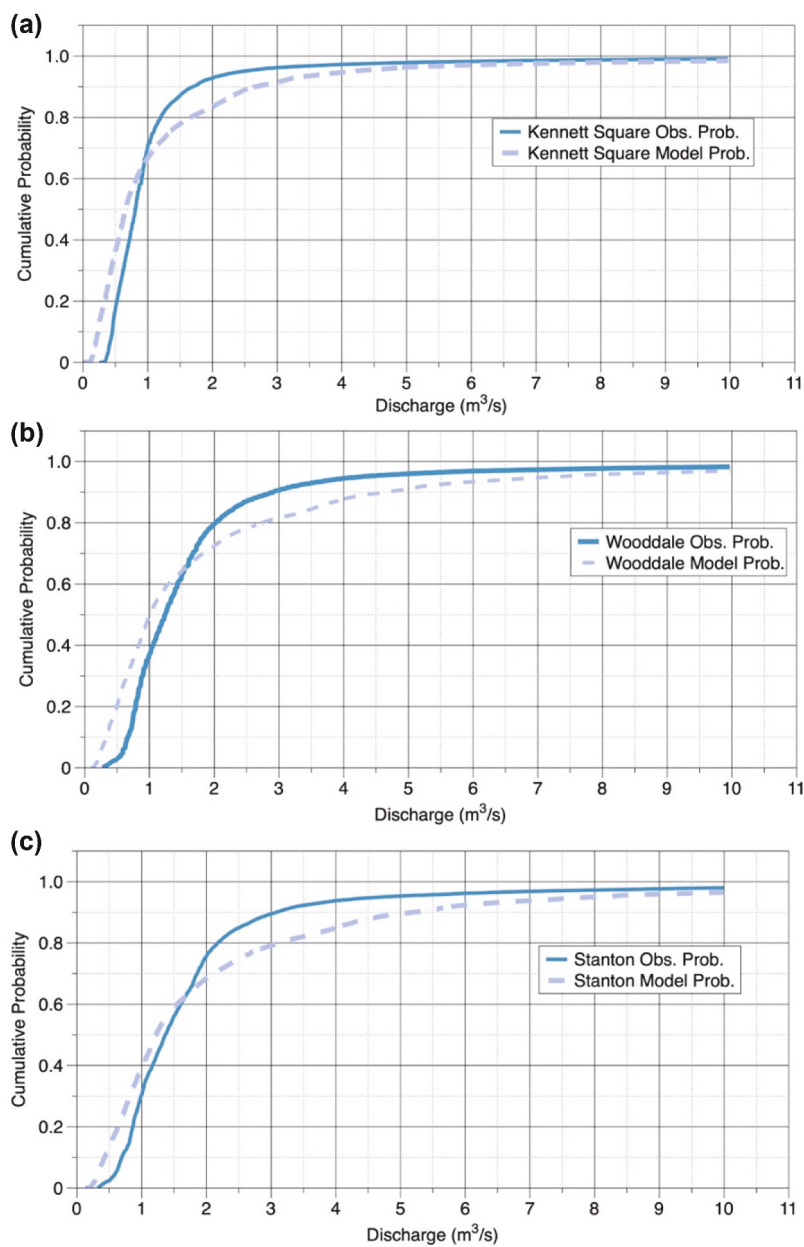


Figure 7. Cumulative probability curves for Kennett Square, PA (a), Wooddale, DE (b) and Stanton, DE (c) for observed (USGS) and modeled (NWM) discharge values. Please note that discharge was limited to 10 m³/s for graphing purposes.

Flooding events

The onset of flooding conditions at Wooddale, DE begin with a discharge of approximately 56 m³/s. Using this definition, five flooding events were recorded at Wooddale from 2016 to 2018. During those five events, the NWM overestimated discharge three

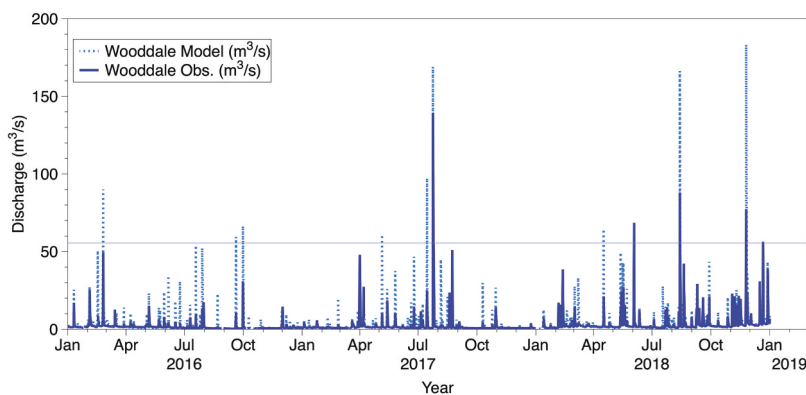


Figure 8. Time series of observed (USGS) and modeled (NWM) discharge for Wooddale, DE (m^3/s). Modeled (dotted) and observations (solid). Horizontal line represents the onset of flood conditions at Wooddale, DE.

Table 4. Observed (USGS) discharge, NWM discharge, difference (NWM-Obs.), precipitation amount at the Bucktoe, PA DEOS station, and the precipitation type for the flooding conditions at Wooddale, DE, where “Conv” is convective, and “MLC” is midlatitude cyclone stratiform.

Date	Obs. Discharge (m^3/s)	NWM Discharge (m^3/s)	Difference (m^3/s)	Precipitation (mm)	Precipitation Type
7/24/17	139	155	16	59.2	Conv.
8/13/18	88	166	78	65.8	Conv.
11/25/18	77	183	106	58.4	MLC
6/3/18	68	4	−64	62.2	Conv.
12/21/18	56	35	−21	58.9	MLC

times and underestimated discharge twice, with a mean absolute error of $57 \text{ m}^3/\text{s}$. [Table 4](#) details the discharge characteristics and meteorological conditions during each event. As expected during flooding events, each was associated with relatively large precipitation totals at the Delaware Environmental Observing System (DEOS) meteorological station at Bucktoe, PA located 10 km from Wooddale. It is worth noting that both convective precipitation events (those which are typically shorter but more intense in character) and mid-latitude cyclone-stratiform precipitation events (typically longer in duration but less intense) are associated with the floods and with both overestimation and underestimation by the NWM.

[Figure 9](#) shows the observed and NWM discharge for each of the three convective flooding events, and [Figure 10](#) shows the same but for flooding associated with the mid-latitude cyclones. Two of the three convective events were associated with overestimation of the discharge by the NWM ([Figure 9a](#) and [c](#)). [Figure 9b](#) shows a large underestimation of the flooding event of 3 June 2018 by the NWM, likely associated with localized convective precipitation. For the flooding events associated with mid-latitude cyclones, the 25 November 2018 event was associated with a large overestimation of discharge by the model ([Figure 10a](#); [Table 4](#)), while the NWM underestimated the discharge during the flooding event of 21 December 2018. For these flooding events, there seems to be

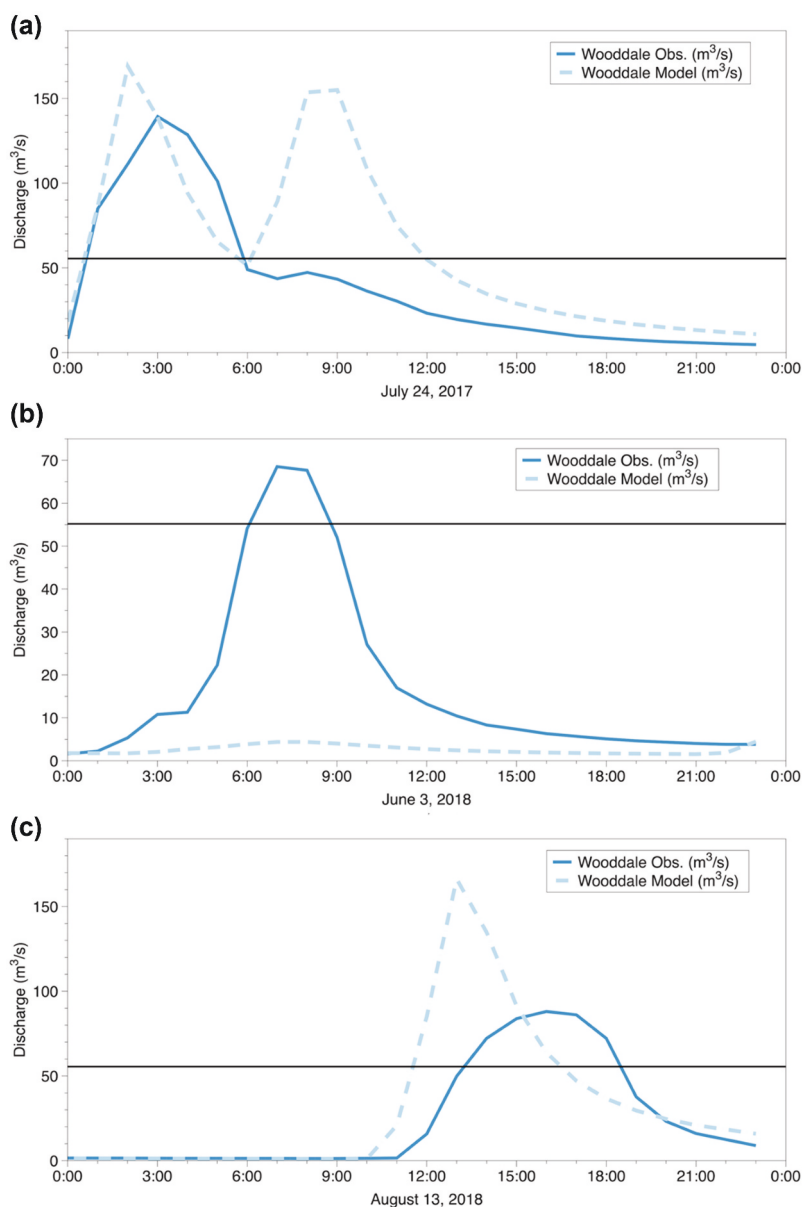


Figure 9. Hydrographs of discharge for Wooddale, PA for all flooding events associated with convective storms; observations (solid) and NWM (dashed). Black line shows the flood stage.

no relationship between the characteristics of the storm (convective or mid-latitude cyclone-stratiform) and the size or sign of the model deviations from observed discharge. In addition, in all cases the NWM peak flow predictions occurred prior to observed peak flow conditions. While this is a small sample size, it does point to issues associated with the NWM when flooding conditions occur along Red Clay Creek.

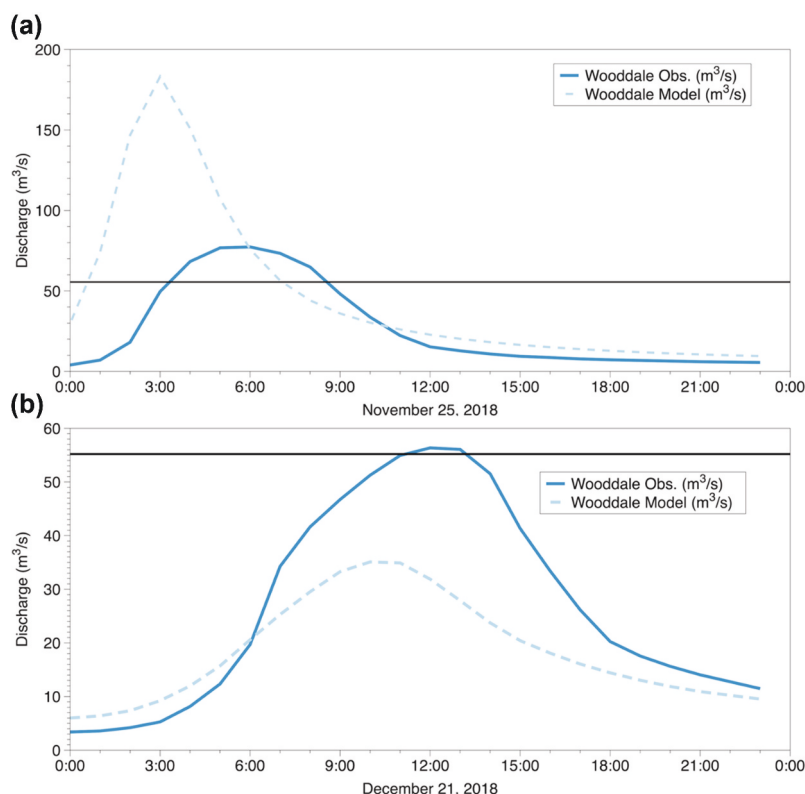


Figure 10. Hydrographs of discharge for Wooddale, PA for all flooding events associated with mid-latitude cyclones; observations (solid) and NWM (dashed). Black line shows the flood stage.

Analysis of extreme NWM versus observation differences

To better understand the situations in which the NWM poorly estimates discharge in the basin, data were collected for Wooddale, DE. There were 60 instances with hourly differences between the NWM and observed discharge of plus or minus 40 m³/s which were retained for analysis. In all, 15 separate meteorological events were responsible for NWM versus observed discharge differences of this magnitude or larger (of the 60 hourly differences all occurred during one of 15 events) with the NWM overestimating the observed values in all but one case. The dates, discharge difference, precipitation total (from the Bucktoe DEOS station) and precipitation type are given for each event in Table 5. Of the 15 events, 10 had precipitation of a convective nature, and two of these were associated with tropical systems moving through the area. The other five events were generally stratiform precipitation associated with the passage of a mid-latitude cyclone through the region. Thus, 66% of the largest differences between the NWM and observed discharge were associated with meso-scale convective precipitation.

Meteorological Analysis

Figure 11 shows the composited sea level pressure, 500 hPa geopotential heights, precipitation rate (mm/day), and lifted index for the eight non-tropical convective events that resulted in large discharge differences between the NWM and observations. The composited sea level

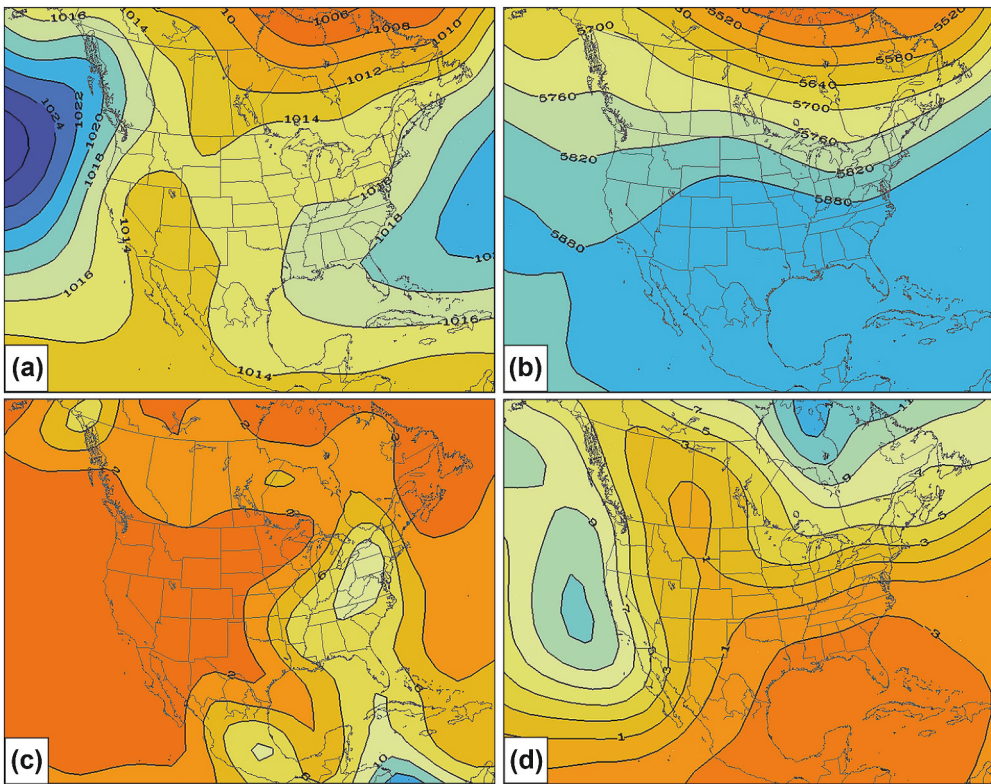


Figure 11. Maps showing a) sea-level pressure, b) 500 hPa heights, c) precipitation rate (mm) and d) lifted index for the eight non-tropical convective events associated with the largest model/observation differences. In all panels, shades of blue indicate higher values and shades of orange indicate lower values.

pressure pattern shows a strong area of high pressure over the Atlantic Ocean producing southwesterly winds and the advection of warm, moist air across the region (Figure 11a). The composited lifted index of this air is less than zero, indicating at least weak instability (Figure 11d). A weak 500 hPa trough to the west of the Red Clay basin helps to provide uplift associated with positive vorticity advection (Figure 11b). The resulting composite precipitation rate (Figure 11c) shows mean values of 8 mm/day. Of course, as in the cases investigated here, much higher amounts of precipitation occurred in isolated convective cells across the region (Table 5). Thus, the convective events, which all occurred from May through August, are associated with relatively heavy convective rainfall resulting from the advection of moist, unstable air into the region via strong return flow on the western side of high pressure over the Atlantic Ocean. Weak uplift results from a 500 hPa trough to the west resulting in heavy convective rainfall.

For the five mid-latitude cyclone events (Figure 12), a strong surface low pressure system is found just to the west of the study region (Figure 12a), in association with a strong 500 hPa trough located over the Ohio Valley (Figure 12b). The transport of moist air from the Atlantic Ocean and uplift associated with quasi-geostrophic processes leads to large mean precipitation rates for the Red Clay Basin of over 10 mm/day (Figure 12c),

Table 5. Date, discharge difference (m^3/s ; model minus observed), precipitation total (mm) from the Bucktoe, PA DEOS station, and general precipitation type of all events where observed and modeled discharge values for Wooddale, DE exceeded plus or minus $40 \text{ m}^3/\text{s}$.

	Delta Discharge (m^3/s)	Precipitation (mm)	Precipitation Type
11/25/2018	133.7	58.4	Mid-Latitude Cyclone
08/13/2018	116.4	65.8	Convective
07/24/2017	111.7	37.9	Convective
07/14/2017	85.8	47.9	Convective
02/25/2016	72.7	33.3	Mid-Latitude Cyclone
09/19/2016	58.8	52.9	Convective - Tropical
05/05/2017	57.7	26.7	Mid-Latitude Cyclone
07/18/2016	52.9	43.0	Convective
07/28/2016	51.3	2.6	Convective
09/30/2016	50.7	48.0	Mid-Latitude Cyclone
04/16/2018	50.6	26.1	Mid-Latitude Cyclone
05/13/2018	44.5	27.0	Convective
06/24/2017	44.3	28.4	Convective - Tropical
08/05/2017	43.3	14.0	Convective
06/03/2018	-64.2	62.2	Convective

with higher values found in localized areas. Thus, the mid-latitude cyclone events are associated with stratiform precipitation resulting from strong upward motion associated with classic quasi-geostrophic dynamics.

Precipitation and soil moisture

“Phase diagrams” were produced for two DEOS meteorological stations located within the basin to further explore the events that produced the largest differences between the NWM and observations. The phase diagrams were constructed by plotting the daily volumetric water content (VWC; percentage) of the upper 20 cm of the soil the day before the event and the precipitation that occurred at each station the day of the event (Figure 13). The 15 discharge events with the largest differences between the NWM and observations are shown as red dots in the phase diagrams. All large differences were associated with rainfall of at least 14 mm (~ 0.5 inches) which was expected (relatively heavy rain must fall to produce a large difference). However, results also show that 13 of 15 large difference events were associated with lower than mean VWC at Bucktoe, PA and 13 of 14 at Mt. Cuba, DE (one precipitation measurement was missing at the Mt. Cuba DEOS station), suggesting that the NWM performed most poorly during heavy rainfall falling on to relatively dry soils.

Summary and Discussion

Modeled discharge values from the National Water Model and USGS stream gage discharge observations were compared for a suburban watershed, the Red Clay Creek in Pennsylvania and Delaware, for 2016–2018. 18-hour retrospective simulations from the NWM were used with contemporaneous hourly USGS discharge observations for three gages located along the stream. These gaging stations included Kennett Square, Pennsylvania (USGS ID: 01479820, NWM Reach ID: 4651090), Wooddale, Delaware (USGS ID: 01479820, NWM Reach ID: 4651912), and Stanton, Delaware (USGS ID: 01480015, NWM Reach ID: 4651930). Major findings from this analysis include the following:

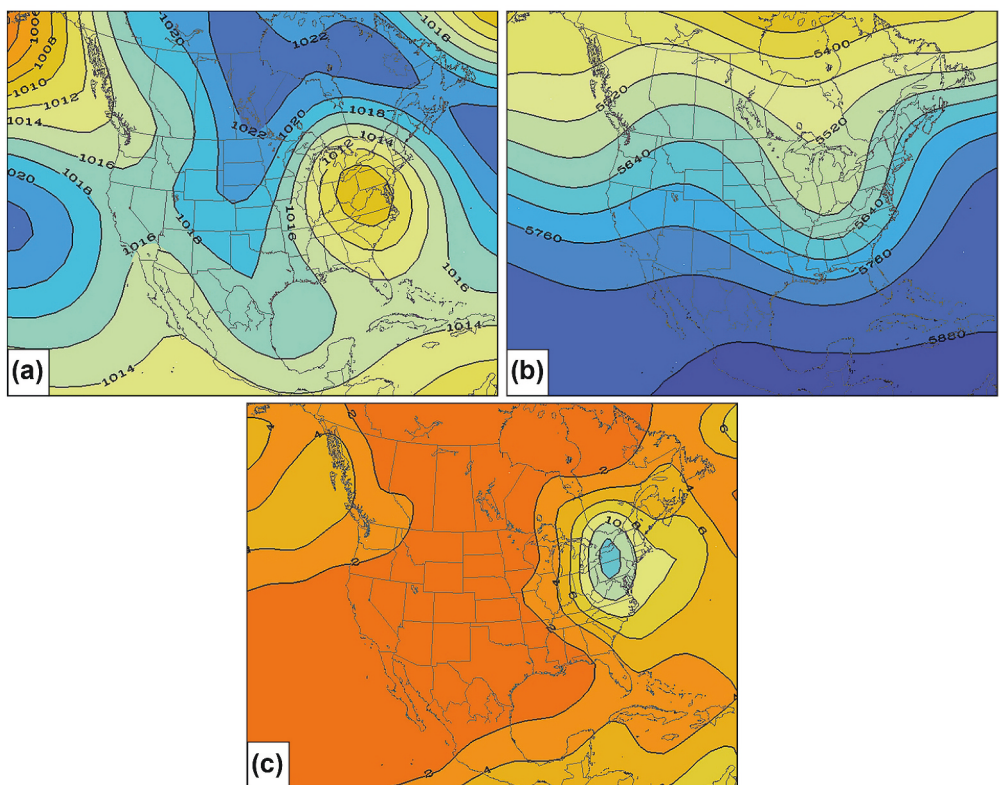


Figure 12. Maps showing a) sea-level pressure, b) 500 hPa heights, and c) precipitation rate (mm) for the five mid-latitude cyclone events associated with the largest model/observation differences. In all panels, shades of blue indicate higher values and shades of orange indicate lower values.

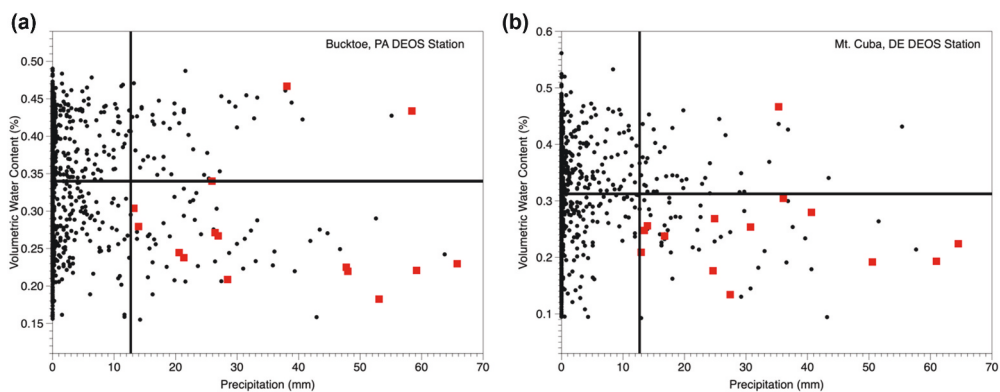


Figure 13. Phase diagrams for the two DEOS mesonet stations within the Red Clay basin a) Bucktoe, PA and b) Mount Cuba, DE. Phase diagrams plot the volumetric water content of the day before the event along the y-axis and precipitation of the day of the event along the x-axis.

- The mean of discharge estimates from the NWM and from USGS observations were found to be significantly different from one another using a variety of standard statistical tests (t-test, Z-test, Wilcoxon signed-rank test).
- The NWM generally underestimates low-flow conditions and overestimates high-flow conditions throughout the Red Clay Creek watershed.
- The relationship between NWM discharge estimates and USGS observations is weakest for the smallest watershed area and improves with an increase in the watershed size.
- Convective rainfall events were associated with 66% of the 15 largest differences between NWM discharge estimates and USGS observations. Mid-latitude cyclone stratiform precipitation events accounted for the other 34%
- Largest differences (13 out of 15 occurrences) between the NWM and observations occurred with pre-cursor soil moisture that was below the mean (dry soil conditions), in conjunction with heavy rainfall.

Taken as a whole, the results of this analysis add to the growing evaluation of the performance of the NWM for varying stream types situated in diverse locations across the United States. In this case, the performance of the NWM was evaluated for a small, suburban watershed in a generally humid area with mixed land use/land cover types. Like previous studies, the analysis here suggests that the NWM underestimates low-flow conditions, potentially associated with ground water exchange processes within the model (Bales & Flowers, 2021; Hansen et al., 2019; Karki et al., 2021). In addition, the accuracy of the model predictions increased as the size of the basin increased within the Red Clay Creek Basin, much as Rojas et al. (2020) found for streams across Iowa. And, like Duan and Kumar (2020), soil moisture was found to be an important variable in those events where the NWM performed most poorly.

The prediction of water levels along Red Clay Creek are important as it serves as a source of water for human consumption and has experienced severe flooding events in the past. Hence, predictions of both high and low-flows have great societal relevance. During five flooding events at Wooddale, DE (from 2016 to 2018) the NWM overestimated discharge three times and underestimated discharge twice. The sign of the model error had no connection to the type of precipitation event (stratiform or convective). From this small sample it is difficult to draw strong conclusions on how the NWM may perform with future flooding events but warrants further investigation. Given the lack of homogeneity in soil moisture measurements across the United States and the fact that soil moisture conditions heavily influence runoff (Baker et al., 2022; Cosh et al., 2021), further analysis on soil's influence on NWM predictions is necessary. As such, future research will evaluate the NWM over extended periods to assess its performance during flooding events in the context of meteorological forcings and pre-cursor soil conditions.

It is important that water resource managers, emergency managers and other decision makers are aware of the tendencies of the NWM as its use becomes more widespread. The Red Clay Creek watershed is like other suburban watersheds across the mid-Atlantic region of the United States as it comprises a mixture of urban and mixed land cover uses (Wardrop et al., 2005). As the NWM continues to undergo validation, the results of this study should be compared to similar geographical and climatological areas to ascertain any tendencies in the model that can provide useful information to emergency managers and water resource

professionals. Studies such as this can aid in the continued improvement of the model, and in a better understanding of its performance in similar regions across the country.

Acknowledgments

The authors would like to thank the anonymous peer reviewers for their suggestions as well as Kevin Brinson and Chris Hughes for their helpful comments whilst revising this manuscript for publication.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This publication was made possible by the National Science Foundation EPSCoR Grant No. 1757353 and the State of Delaware.

ORCID

Saber E. Brasher  <http://orcid.org/0000-0003-2925-5691>

Daniel J. Leathers  <http://orcid.org/0000-0001-9002-9031>

Data availability statement

The authors confirm that data used for this research are publicly available and can be accessed by the links and references provided.

References

- Baker, C. B., Cosh, M., Bolten, J., Brusberg, M., Caldwell, T., Connolly, S., Dobrev, I., Edwards, N., Goble, P. E., Ochsner, T. E., Quiring, S. M., Robotham, M., Skumanich, M., Svoboda, M., White, W. A., & Woloszyn, M. (2022). Working toward a National coordinated soil moisture Monitoring network: Vision, progress, and future directions. *Bulletin of the American Meteorological Society*, 103 (12), E2719–E2732. <https://doi.org/10.1175/BAMS-D-21-0178.1>
- Bales, J., & Flowers, T. (2021). Featured collection introduction: National water Model III. *JAWRA Journal of the American Water Resources Association*, 57(2), 205–208. <https://doi.org/10.1111/1752-1688.12913>
- Cosh, M. H., Caldwell, T. G., Baker, C. B., Bolten, J. D., Edwards, N., Goble, P., Hofman, H., Ochsner, T. E., Quiring, S., Schalk, C., Skumanich, M., Svoboda, M., & Woloszyn, M. E. (2021). Developing a strategy for the national coordinated soil moisture monitoring network. *Vadose Zone Journal*, 20(4). <https://doi.org/10.1002/vzj2.20139>
- Cruz-Ortiz, C., & Miller, K. (2013). Economic value of the red clay creek watershed, Accessed September 20. https://www.wrc.udel.edu/wp-content/uploads/2016/10/Economic-Value-of-the-Red-Clay-Creek_DRAFT.pdf
- Davis, J. (1986). *Statistics and Data Science in Geology* (2nd ed.). Wiley.
- Degaetano, A. T. (2009). Time-dependent changes in extreme-precipitation return-period amounts in the continental United states. *Journal of Applied Meteorology and Climatology*, 48 (10), 2086–2099. <https://doi.org/10.1175/2009JAMC2179.1>

- Dieter, C. A., Maupin, M. A., Caldwell, R. R., Harris, M. A., Ivahnenko, T. I., Lovelace, J. K., Barber, N. L., & Linsey, K. S. (2018). Estimated use of water in the United States in 2015: U.S. Geological Survey Circular, 1441. [Supersedes USGS Open-File Report 2017–1131.]. <https://doi.org/10.3133/cir1441>
- Duan, Y., Akula, S., Kumar, S., Lee, W., & Khajehei, S. (2023). A hybrid physics–AI model to improve hydrological forecasts. *Artificial Intelligence for the Earth Systems*, 2(1). <https://doi.org/10.1175/AIES-D-22>
- Duan, Y., & Kumar, S. (2020). Predictability of seasonal streamflow and soil moisture in National water model and a humid Alabama–coosa–Tallapoosa River basin. *Journal of Hydrometeorology*, 21(7), 1447–1467. <https://doi.org/10.1175/JHM-D-19>
- Griffiths, M. L., & Bradley, R. S. (2007). Variations of twentieth-century temperature and precipitation extreme indicators in the northeast United States. *Journal of Climate*, 20(21), 5401–5417. <https://doi.org/10.1175/2007JCLI1594.1>
- Hammond, J. C., Doheny, E. J., Dillow, J. J. A., Nardi, M. R., Steeves, P. A., & Warner, D. L. (2022). *Peak-flow and low-flow magnitude estimates at defined frequencies and durations for nontidal streams in Delaware: U.S. Geological Survey Scientific Investigations Report 2022–5005* (p. 46). <https://doi.org/10.3133/sir20225005>
- Hansen, C., Shafiei Shiva, J., McDonald, S., & Nabors, A. (2019). Assessing retrospective National water model streamflow with respect to droughts and low flows in the Colorado River basin. *JAWRA Journal of the American Water Resources Association*, 55(4), 964–975. <https://doi.org/10.1111/1752-1688.12784>
- Jachens, E. R., Hutcheson, H., Thomas, M. B., & Steward, D. R. (2021). Effects of Groundwater-surface water exchange Mechanism in the National water model over the northern high Plains aquifer, USA. *JAWRA Journal of the American Water Resources Association*, 57(2), 241–255. <https://doi.org/10.1111/1752-1688.12869>
- Karki, R., Krienert, J. M., Hong, M., & Steward, D. R. (2021). Evaluating baseflow simulation in the National water model: A case study in the northern high Plains region, USA. *JAWRA Journal of the American Water Resources Association*, 57(2), 267–280. <https://doi.org/10.1111/1752-1688.12911>
- Kauffman, G. J., Belden, A. C., Vonck, K. J., & Homsey, A. R. (2009). Link between impervious cover and base flow in the White Clay Creek Wild and scenic watershed in Delaware. *Journal of Hydrologic Engineering*, 14(4), 324–334. [https://doi.org/10.1061/\(asce\)1084-0699\(2009\)14:4\(324\)](https://doi.org/10.1061/(asce)1084-0699(2009)14:4(324))
- Kunkel, K. E., Karl, T. R., Brooks, H., Kossin, J., Lawrimore, J. H., Arndt, D., Bosart, L., Changnon, D., Cutter, S. L., Doesken, N., Emanuel, K., Groisman, P. Y., Katz, R. W., Knutson, T., O'brien, J., Paciorek, C. J., Peterson, T. C., Redmond, K., Robinson, D. ... Wuebbles, D. (2013). Monitoring and understanding trends in extreme storms: State of knowledge. *Bulletin of the American Meteorological Society*, 94(4), 499–514. <https://doi.org/10.1175/BAMS-D-11-00262.1>
- Leathers, D. J., Brasher, S. E., Brinson, K. R., Hughes, C., & Weiskopf, S. (2020). A comparison of extreme precipitation event frequency and magnitude using a high-resolution rain gage network and NOAA atlas 14 across Delaware. *International Journal of Climatology*, 40(8), 3748–3756. <https://doi.org/10.1002/joc.6425>
- Moltz, H. L. N., Palmer, J. B., & Smith, Z. M. (2018). Streamflow alteration from impervious cover: Are all watersheds created equal? *JAWRA Journal of the American Water Resources Association*, 54(6), 1222–1238. <https://doi.org/10.1111/1752-1688.12681>
- NOAA National Water Model CONUS Retrospective Dataset was Retrieved June 23, 2023 from <https://registry.opendata.aws/nwm-archive> .
- Rojas, M., Quintero, F., & Krajewski, W. F. (2020). Performance of the National water model in Iowa using Independent observations. *JAWRA Journal of the American Water Resources Association*, 56(4), 568–585. <https://doi.org/10.1111/1752-1688.12820>
- Senior, L. A., & Koerkle, E. H. (2003). 'Simulation of streamflow and water quality in the Red Clay Creek subbasin of the Christina River basin, Pennsylvania and Delaware, 1994–98: U.S. Geological Survey Water-Resources Investigations Report', 2003–4138, 119. <https://pubs.er.usgs.gov/publication/wri034138>

- Shuster, W. D., Bonta, J., Thurston, H., Warnemuende, E., & Smith, D. R. (2005). Impacts of impervious surface on watershed hydrology: A review. *Urban Water Journal*, 2(4), 263–275. <https://doi.org/10.1080/15730620500386529>
- Smith, B. K., & Smith, J. A. (2015). The flashiest watersheds in the contiguous United States. *Journal of Hydrometeorology*, 16(6), 2365–2381. <https://doi.org/10.1175/JHM-D-14-0217.1>
- Trenberth, K. E. (2011). Changes in precipitation with climate change. *Climate Research*, 47(1–2), 123–138. <https://doi.org/10.3354/cr00953>
- Turnipseed, D. P., & Sauer, V. B. (2010). Discharge measurements at gaging stations: U.S. Geological Survey techniques and methods. *Book*, 3, A8, 87. (Also available at). <https://pubs.usgs.gov/tm/tm3-a8/>)
- U.S. Geological Survey, 2023, National water information System data available on the World Wide web (water data for the nation), Retrieved June 23, 2023 <http://waterdata.usgs.gov/nwis/> .
- Viterbo, F., Mahoney, K., Read, L., Salas, F., Bates, B., Elliott, J., Cosgrove, B., Dugger, A., Gochis, D. & Cifelli, R. (2020). A multiscale, hydrometeorological forecast evaluation of national water model forecasts of the May 2018 Ellicott City, Maryland, flood. *Journal of Hydrometeorology*, 21(3), 475–499. <https://doi.org/10.1175/JHM-D-19-0125.1>
- Viterbo, F., Read, L., Nowak, K., Wood, A. W., Gochis, D., Cifelli, R., & Hughes, M. (2020). General Assessment of the operational Utility of National water model reservoir inflows for the Bureau of Reclamation Facilities. *Water*, 12(10), 2897. <https://doi.org/10.3390/w12102897>
- Wardrop, D. H., Bishop, J. A., Easterling, M., Hychka, K., Myers, W., Patil, G. P., & Taillie, C. (2005). Use of landscape and land use parameters for classification and characterization of watersheds in the mid-Atlantic across five physiographic provinces. *Environmental and Ecological Statistics*, 12(2), 209–223. <https://doi.org/10.1007/s10651-005-1042-5>