How Do Satellite Precipitation Products Affect Water Quality Simulations? A Comparative Analysis of Rainfall Datasets for River Flow and Riverine Nitrate Load in an Agricultural Watershed

Mahesh R. Tapas [1,2] (ID)

1   Integrated Coastal Program, Department of Coastal Studies, East Carolina University,
    Greenville, NC 27858, USA; tapas.4@osu.edu
2   Department of Extension and Department of Food, Agricultural and Biological Engineering, The Ohio State
    University, Columbus, OH 43210, USA

Abstract: Excessive nitrate loading from agricultural runoff leads to substantial environmental and economic harm, and although hydrological models are used to mitigate these effects, the influence of various satellite precipitation products (SPPs) on nitrate load simulations is often overlooked. This study addresses this research gap by evaluating the impacts of using different satellite precipitation products—ERA5, IMERG, and gridMET—on flow and nitrate load simulations with the Soil and Water Assessment Tool Plus (SWAT+), using the Tar-Pamlico watershed as a case study. Although agricultural activities are higher in the summer, this study found the lowest nitrate load during this season due to reduced runoff. In contrast, the nitrate load was higher in the winter because of increased runoff, highlighting the dominance of water flow in driving riverine nitrate load. This study found that although IMERG predicts the highest annual average flow (120 m^3/s in Pamlico Sound), it unexpectedly results in the lowest annual average nitrate load (1750 metric tons/year). In contrast, gridMET estimates significantly higher annual average nitrate loads (3850 metric tons/year). This discrepancy underscores the crucial impact of rainfall datasets on nitrate transport predictions and highlights how the choice of dataset can significantly influence nitrate load simulations.

Keywords: rainfall datasets; hydrological modeling; nitrate; coastal watershed; water quality; discharge; agriculture; simulation; eutrophication; Soil and Water Assessment Tool Plus (SWAT+)

1. Introduction

Understanding hydrological and nutrient dynamics is vital for effective environmental management and sustainable water resources, particularly in coastal watersheds vulnerable to nutrient pollution [1–4]. Accurate modeling of rainfall and nitrate transport is crucial for predicting water flow and nutrient movement through river systems, helping maintain ecosystem health. Nitrate runoff, often driven by agriculture and urbanization, leads to issues like eutrophication, algal blooms, and hypoxia, which degrade water quality and threaten biodiversity [5–8]. Reliable rainfall datasets are essential for precise hydrological modeling [9], as they directly influence assessments of flow variations and nutrient loading across watersheds. Enhanced models can inform water resource management strategies, guiding efforts to mitigate pollution, protect ecosystems, and ensure long-term water sustainability [1,9].

Tapas et al., 2024a [10] highlight the importance of updated policies for the Tar-Pamlico watershed, emphasizing its vulnerability to nitrate runoff and the need for adaptive management strategies. The Tar-Pamlico basin, a major coastal watershed in eastern North Carolina, drains into the Pamlico Sound, the largest U.S. lagoon on the east coast. Covering over 6400 square miles, the basin spans from the upper Piedmont to the coastal plain, encompassing diverse land uses—agricultural, urban, and forest [10–15]. Ecologically, it supports rich habitats and species [10], while serving as a critical water source for

communities and industries [16]. Due to its susceptibility to nitrate runoff, the Tar-Pamlico basin is a significant case study for understanding the interaction between rainfall, runoff, and nutrient transport [10]. This complexity makes it ideal for investigating hydrological dynamics and informing water management strategies [10,17].

Nutrient pollution, particularly from nitrate runoff caused by agricultural activities, urban development, and wastewater discharge, is a major challenge in watersheds like the Tar-Pamlico. Excess nitrate can trigger algal blooms, which deplete oxygen where aquatic life cannot thrive [10,17]. This eutrophication process not only disrupts ecosystems but also threatens drinking water supplies, biodiversity, and recreational and commercial fishing [18]. Accurate modeling of nitrate transport is therefore critical for predicting these impacts and developing effective pollution mitigation strategies [10].

Rainfall plays a central role in hydrological studies, as it drives surface runoff and influences the movement of water and nutrients through river systems [19]. High-quality rainfall datasets are essential for modeling these processes, helping predict flow fluctuations, nutrient loading, and the risk of flooding or drought [20]. Inconsistent rainfall data can result in flawed models, undermining predictions of water availability and nutrient pollution, which in turn hampers effective water resource management and ecosystem protection [21].

This study compares three prominent rainfall datasets—ERA5, IMERG, and gridMET—each essential for hydrological and environmental modeling. ERA5 is a global reanalysis dataset offering high-resolution climate and hydrological data [9]. IMERG, part of the NASA GPM mission, provides high-resolution satellite-based precipitation data, useful for capturing extreme weather events [22]. GridMET focuses on the continental U.S., offering high-resolution data for studies related to evapotranspiration, drought, and agricultural impacts [23]. Tapas et al. (2024) found that gridMET had the highest correlation with observed rainfall, followed by IMERG, while ERA5 performed weaker at daily scales [9]. This study extends the analysis to a monthly scale to assess how these datasets influence flow and nitrate transport predictions, providing insights into their broader hydrological modeling performance.

Current rainfall datasets exhibit gaps and inconsistencies that can affect hydrological models, especially for monthly scale nitrate transport predictions. Satellite datasets like IMERG may capture short-term rainfall events well but can overestimate monthly totals, inflating predictions of flow and nitrate runoff [24]. In contrast, reanalysis of datasets like ERA5 may underrepresent localized rainfall patterns, leading to underestimates of flow and nutrient loads [25]. Addressing these inconsistencies is critical for improving the accuracy of hydrological models and refining nutrient management strategies.

This study aims to enhance our understanding of hydrological and nutrient dynamics by comparing the performance of ERA5, IMERG, and gridMET in predicting flow and nitrate transport in the Tar-Pamlico basin. Specifically, it analyzes seasonal rainfall estimates from these datasets and their influence on flow variation and nitrate loading across five sub-watersheds, including the Upper Tar, Fishing Creek, Lower Tar, Pamlico, and Pamlico Sound. By exploring the impact of different rainfall inputs on the flow–nitrate relationship, this study offers valuable insights into the spatial and temporal distribution of nitrate loads, which can improve hydrological modeling and inform water management strategies.

Understanding the flow–nitrate relationship is crucial for predicting nutrient loading and its environmental impacts [10]. Accurate models enable researchers to identify critical periods of nutrient runoff, which often coincide with seasonal rainfall and agricultural practices. By improving these models, especially in capturing flow dynamics across various sub-watersheds, scientists can make more precise predictions about nitrate loading into downstream ecosystems. This improved understanding is essential for developing effective water management strategies to reduce nutrient pollution, prevent algal blooms and hypoxia, and protect water quality. Improved models can help policymakers and environmental managers devise more targeted approaches to nutrient management, conservation,

and sustainable agricultural practices, ensuring healthier ecosystems and resilient water resources [1,10].

2. Materials and Methods

2.1. Study Area

The Tar-Pamlico watershed, located in eastern North Carolina (Figure 1), spans over 6400 square miles and is a significant hydrological region that extends from the hills of the Piedmont to the flat, flood-prone coastal plain [10,26]. The Tar River originates in the Piedmont region of North Carolina, is named after the historical tar industry in the area and changes its name to the Pamlico River as it widens near Washington, NC, due to its transition from a freshwater river to a tidal estuary influenced by the Pamlico Sound [10,26,27]. The region's diverse land uses—agricultural, urban, and forested— and current issues with excessive nitrate loadings [10,15] make it a key area for studying the impacts of nitrate pollution, largely driven by agricultural runoff [10]. Nitrate transport, exacerbated by heavy rainfall and flooding, poses a threat to water quality and contributes to eutrophication in the Pamlico Sound. With its varied geography, including elevated, clay-rich soils in the Piedmont and permeable sandy soils in the coastal plain, the watershed exhibits complex hydrological processes, making it an ideal setting for examining the interplay between rainfall, flow, and nutrient loading [10,26].
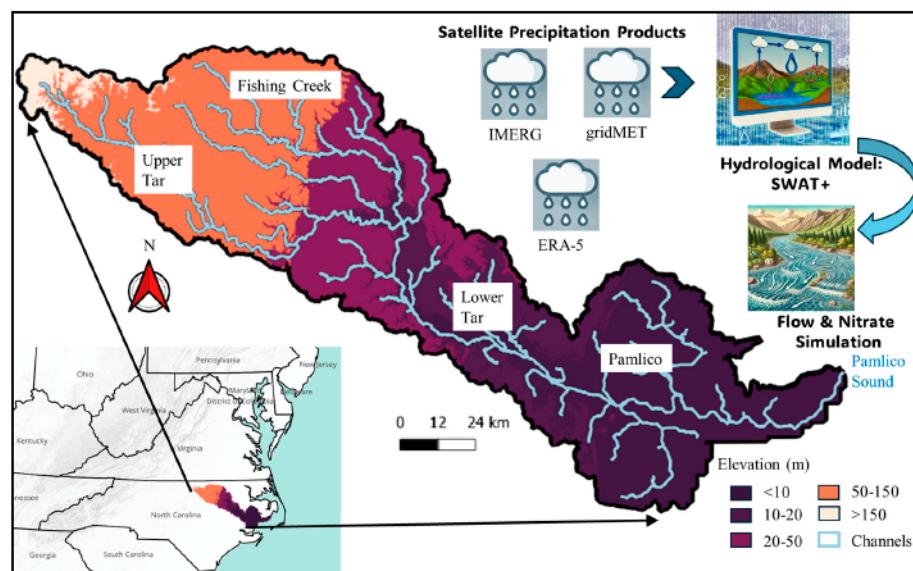
Figure 1. Elevation map of the study area watershed in eastern North Carolina, USA, showing elevation categories from <10 m to >150 m. The inset map provides the geographic location of the watershed within the broader southeastern U.S. region.

2.2. Rainfall Datasets

The study compares three prominent rainfall datasets—ERA5, IMERG, and gridMET— to evaluate their performance in hydrological modeling and rainfall analysis [9][1]. ERA5, a global reanalysis product from the European Centre for Medium-Range Weather Forecasts (ECMWF), provides high-resolution hourly data derived from a 4D-var integration approach, covering atmospheric, land, and sea-state parameters [28]. GridMET, specifically tailored for the continental U.S., merges PRISM's spatial detail with NLDAS-2's temporal precision to deliver daily meteorological data, including precipitation, at a 0.04° resolution. IMERG, part of NASA's Global Precipitation Measurement (GPM) mission, offers high-resolution precipitation estimates by combining data from passive microwave sensors, infrared sensors, and radar with a 0.1° spatial and 30-min temporal resolution [9]. This study adopts a novel approach by investigating how the selection of the above-mentioned SPPs influences riverine nitrate transport simulations in a coastal agricultural watershed,

providing critical insights into the impact of rainfall variability on hydrological flow and nutrient dynamics.

2.3. Hydrological Model

This study utilized the SWAT+ model for the Tar-Pamlico River Basin, developed by Tapas et al., 2024a, 2024b [10,15], and simulated with the above-mentioned SPPs [9]—ERA5, IMERG, and gridMET. Tapas et al. [10] developed and optimized SWAT+ (v2.3.3) to simulate hydrological processes, including flow dynamics and nitrate transport, within the Tar-Pamlico basin (Figure 2). The model incorporated comprehensive environmental data, including elevation (USGS), land cover (NLCD), soil data (SSURGO), rainfall (IMERG), temperature, wastewater treatment plant data, and agricultural land use-land management data. The model was rigorously calibrated to optimize monthly flow and nitrate loads, with a focus on agricultural nitrate loss, a limiting nutrient for the Tar-Pamlico watershed. Additionally, the model was enhanced through soft-calibration calibration, accurately simulating crop yields, denitrification, and nitrate export at the Hydrological Response Unit (HRU) level. Cross-validation for monthly flow across multiple locations confirmed the model's robustness, making it an effective tool for evaluating how varying rainfall datasets influence nitrate transport and water management in the Tar-Pamlico watershed [10].
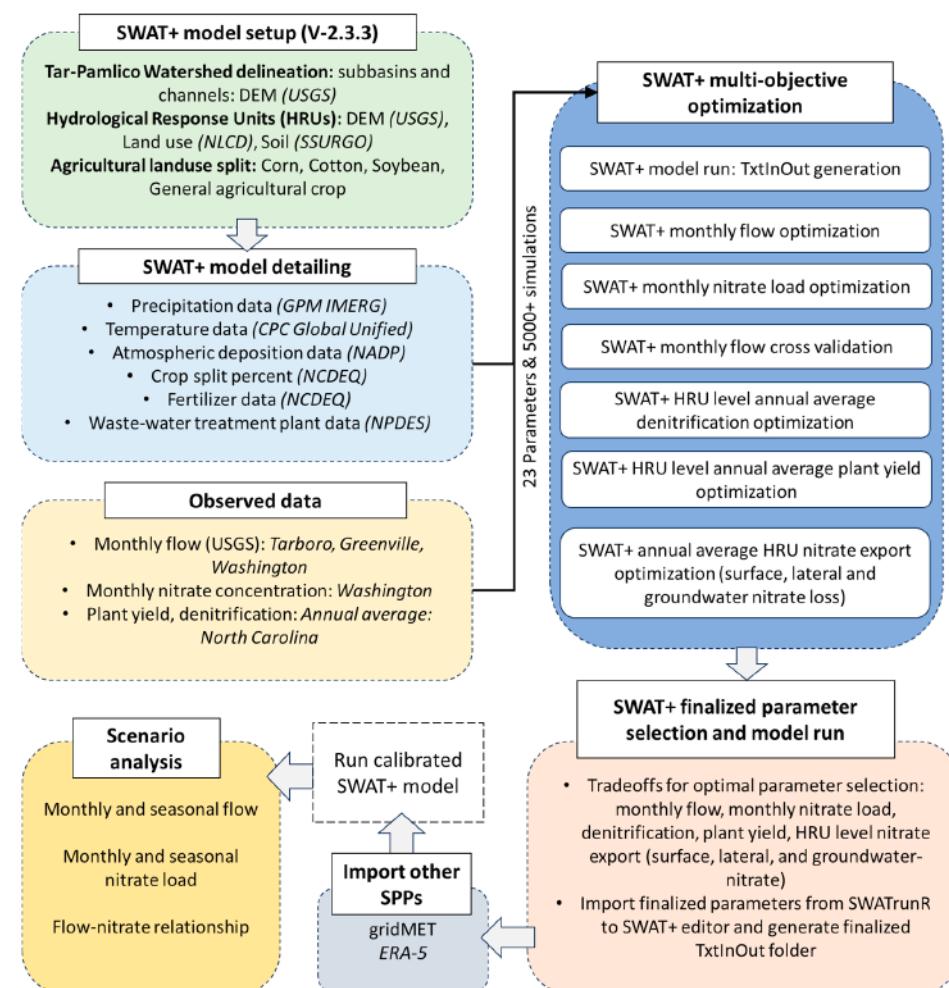
Figure 2. Framework for SWAT+ model setup, calibration, and multi-objective optimization for analyzing flow and nitrate load in the Tar-Pamlico watershed. The diagram outlines key data inputs, model detailing, and scenario analysis with multiple rainfall datasets.

2.4. Flow–Nitrate Relationship Analysis Across Subbasins and Seasons

This study examines the impact of different rainfall datasets on the flow–nitrate dynamics across five subbasins (HUC-8) within the Tar-Pamlico River Basin, including Upper Tar, Fishing Creek, Lower Tar, Pamlico, and Pamlico Sound (Figure S1, Supplementary Information). Using the SWAT+ model, monthly flow and nitrate transport are simulated under varying rainfall inputs from ERA5, IMERG, and gridMET datasets. The analysis focuses on key seasons, such as spring, which is characterized by high flow and nitrate loads due to increased precipitation and agricultural activity, and summer, where reduced rainfall leads to lower flow and nitrate transport. By comparing the flow–nitrate relationship across these subbasins, the study evaluates how different rainfall datasets influence nutrient transport and helps refine strategies for managing nitrate pollution in coastal watersheds[2].

3. Results and Discussion

3.1. Overview of Rainfall Datasets (ERA5, IMERG, and gridMET)

Figure 3 illustrates the average seasonal rainfall data for Greenville, NC, derived from three SPPs—ERA5, gridMET, and IMERG [9]. These datasets have the following distinct characteristics: ERA5 is a high-resolution reanalysis dataset from satellite and in situ observations, gridMET combines meteorological and satellite inputs for the U.S., and IMERG, from the GPM mission, provides near-global, high-resolution satellite-based precipitation estimates [9]. The seasonal analysis, based on average values from January 2001 to December 2019, shows that in the fall, ERA5 reports 320 mm of rainfall, gridMET records 336 mm, and IMERG observes 326 mm. For spring, ERA5 reports 282 mm, gridMET 305 mm, and IMERG 316 mm. Summer exhibits the highest rainfall across all datasets, with ERA5 recording 392 mm, gridMET 428 mm, and IMERG 443 mm (Figure 3). In winter, ERA5 reports 257 mm, gridMET 256 mm, and IMERG 285 mm. This comparison highlights that the summer season has the highest rainfall, and the datasets show consistent agreement, especially during winter, reflecting the robustness of multi-source rainfall measurements. North Carolina experiences the highest rainfall in summer due to frequent thunderstorms, convective activity, and occasional tropical storms from the Atlantic hurricane season [29]. In spring and fall, rainfall is moderate, influenced by frontal systems and seasonal transitions. Winter has the lowest rainfall, as colder temperatures limit moisture availability and reduce the occurrence of heavy precipitation [30]. It is worth noting that Tapas et al. [9] found that gridMET performed the best in accurately representing rainfall patterns in Greenville when compared with the observed rainfall data, followed closely by IMERG, while ERA5 exhibited the weakest performance. However, in an extensive analysis of rainfall datasets for flow simulation, Tapas et al. [9] determined that IMERG performed the best in simulating daily flow, followed by ERA5 and gridMET, when compared to USGS flow station data (02084000) at Greenville, NC.
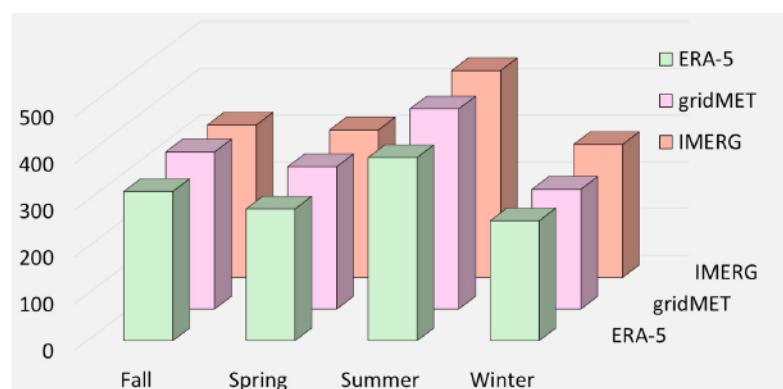

Figure 3. Seasonal rainfall comparison for Greenville, NC (2001–2019) using ERA5, gridMET, and IMERG datasets [The bar chart shows seasonal variations in rainfall, highlighting differences between the three datasets for Fall, Spring, Summer, and Winter].

3.2. Flow Variation Across Locations and Rainfall Datasets

The annual average flow values (Figure 4) from January 2003 to December 2019 reveal that IMERG consistently predicts the highest flow rates across all locations, with Pamlico Sound showing 120 m^3/s, Pamlico 112 m^3/s, and Lower Tar 70 m^3/s[3]. ERA5 generally predicts moderate flows, while gridMET tends to estimate the lowest values, particularly in the Upper Tar, where it records only 18 m^3/s compared to IMERG's 27 m^3/s. This trend highlights the variability in flow predictions based on the rainfall dataset used, which is critical for accurate hydrological modeling and water resource management in the Tar-Pamlico watershed. Additionally, since SWAT+ is a one-dimensional flow model that does not account for backflow, this study is less confident about the results at Pamlico Sound, where such dynamics could significantly influence flow behavior. The complex hydrodynamics of estuarine environments, such as Pamlico Sound, often involve bidirectional flows driven by tides, winds, and other coastal processes, which are not captured by SWAT+ [10]. As a result, the simplified flow assumptions of SWAT+ may lead to inaccuracies in these areas. To improve confidence in flow predictions, especially in regions prone to backflows and other dynamic water movements, it may be necessary to rely on more advanced hydrodynamic models that can simulate multidimensional flow patterns. These models would provide a more comprehensive understanding of water exchanges and interactions, particularly in coastal or estuarine environments like Pamlico Sound [13].
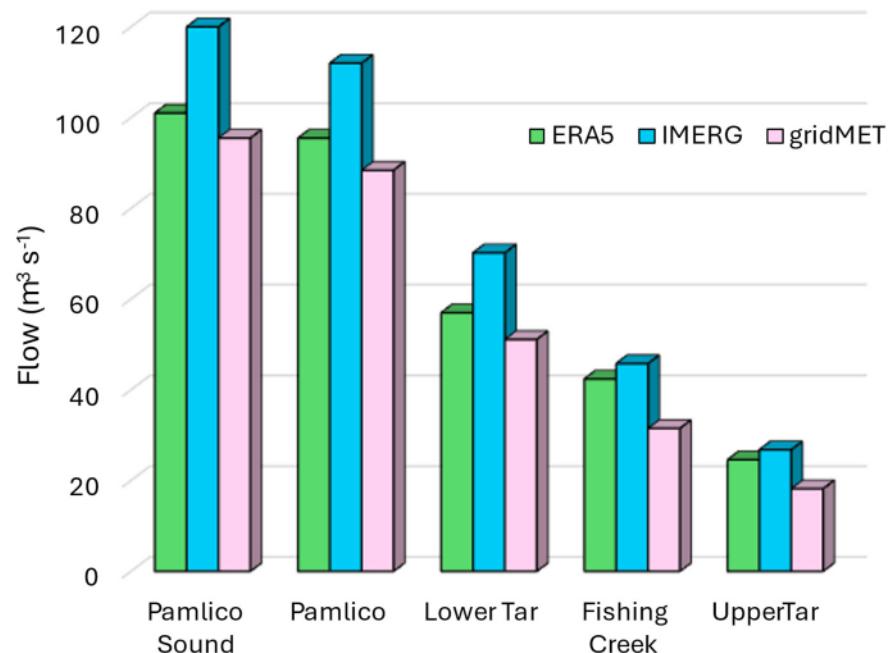
Figure 4. Annual average flow values (m^3/s) from 2003 to 2019 at five locations in the Tar-Pamlico watershed, comparing predictions from three rainfall datasets: ERA5, IMERG, and gridMET [IMERG consistently predicts higher flows across all locations, with gridMET generally estimating the lowest values]. As SWAT+ does not account for backflows, the flow values at Pamlico and Pamlico Sound may be overestimated compared to actual conditions, where backflow could reduce overall flow rates.

The flow maps (Figure 5) for ERA5, gridMET, and IMERG show a consistent pattern of increasing flow downstream, particularly in the Pamlico and Pamlico Sound regions. IMERG predicts the highest flow rates across the watershed, especially in the downstream areas, while ERA5 and gridMET show lower flow estimates, particularly in upstream subbasins. Each location shows distinct flow patterns, with variances across both geography and rainfall datasets, which are important for understanding regional hydrology and water resource management.
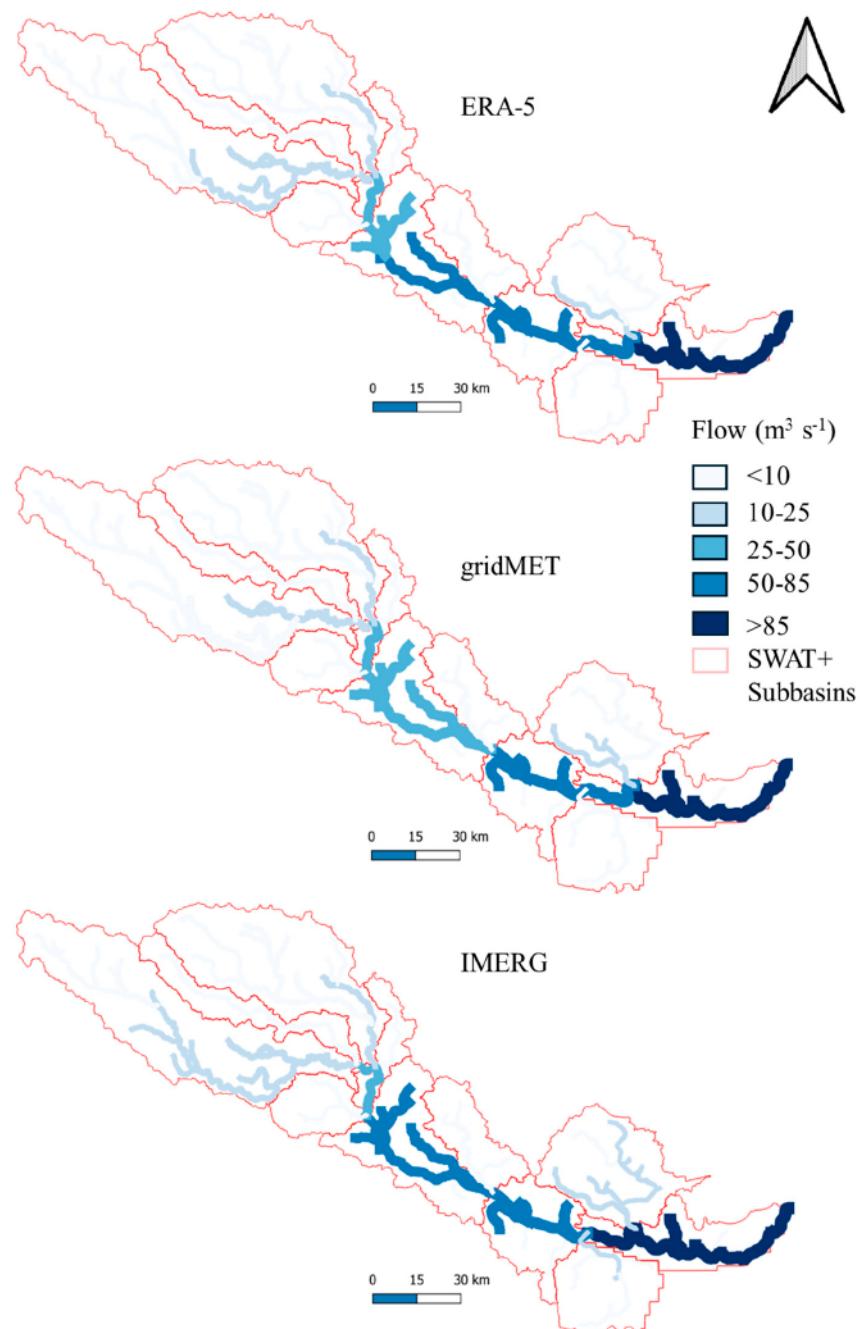
Figure 5. Streamflow comparison for the study area using ERA5, gridMET, and IMERG datasets [The maps depict spatial variations in streamflow (m^3/s) across subbasins, with flow categorized into five classes, highlighting differences in streamflow estimates among the datasets].

This research analyzed the variation in seasonal flow across five regions in the Tar-Pamlico watershed—Upper Tar, Fishing Creek, Lower Tar, Pamlico, and Pamlico Sound—using the IMERG rainfall dataset, as the model was calibrated with this dataset to minimize the number of variables. This study used the most downstream SWAT+ channels to analyze flow values from each corresponding HUC-8 sub-watershed[4]. Flow variation across these locations reveals distinct spatial patterns, with downstream areas like Pamlico Sound consistently exhibiting the highest flow values, while upstream regions such as the Upper Tar and Fishing Creek show significantly lower flow rates.

Seasonal flow variations are key to understanding hydrological dynamics, particularly nitrate loading [10]; even with summer's high rainfall, flow rates peak in Winter and Spring,

when cooler temperatures and lower evapotranspiration allow more water to move through the system. For example, flow in the Upper Tar River reaches 38 m^3/s in Winter, while Pamlico Sound peaks at 155 m^3/s. Higher flow rates during these seasons might elevate nutrient runoff, especially nitrate, posing a risk for eutrophication downstream [10,15]. Spring follows closely, with flows driven by late winter and early spring rains. Pamlico averages 119 m^3/s, while Fishing Creek averages 64 m^3/s. Spring flows coincide with agricultural runoff, increasing nitrate levels and pollution risks.

Summer, despite the highest rainfall, experiences the lowest flows due to high evapotranspiration. During Summer, the flow at the Upper Tar River drops to 13 m^3/s, and Pamlico Sound to 89 m^3/s. These low flows limit nutrient transport but can increase local concentrations, highlighting the need for careful water quality management [31]. Fall sees a moderate flow recovery, with the Lower Tar averaging 57 m^3/s, as consistent rainfall replenishes the watershed post-summer; though lower than Winter or Spring, Fall flows are essential for recharging aquatic systems.

Geographically, Pamlico Sound consistently exhibits the highest flow, serving as the watershed's final receiving body, peaking at 155 m^3/s during Winter. In contrast, upstream regions like Upper Tar and Fishing Creek maintain lower flow rates throughout the year due to smaller drainage areas. This analysis emphasizes the importance of seasonal flow patterns in hydrological models for predicting nutrient transport and addressing water quality issues in the Tar-Pamlico basin [10,31][5].

Figure 6 represents the seasonal analysis of flow data in Washington, NC, for the three rainfall datasets (ERA5, IMERG, and gridMET), revealing clear seasonal trends and variations across the datasets. A time series plot of the datasets illustrates the seasonal dynamics (Figure 6), with IMERG depicted in blue, clearly showing its tendency to yield higher flow rates. Meanwhile, ERA5 and gridMET follow different patterns but with noticeable variability across the seasons. Typically, the higher flow values occur during spring and winter, while the lower values are observed during fall and summer. Interestingly, IMERG appears to consistently provide higher flow values compared to ERA5 and gridMET, particularly during peak flow seasons like spring and winter. This pattern suggests that the model with IMERG data might be more sensitive to heavy rainfall events or that it captures more intense precipitation data, contributing to elevated flow rates. On the other hand, gridMET, while generally lower than ERA5 and IMERG, tends to stabilize the flow values, showing a more conservative estimate that might reflect a model focused on long-term average rainfall (Figure 6).
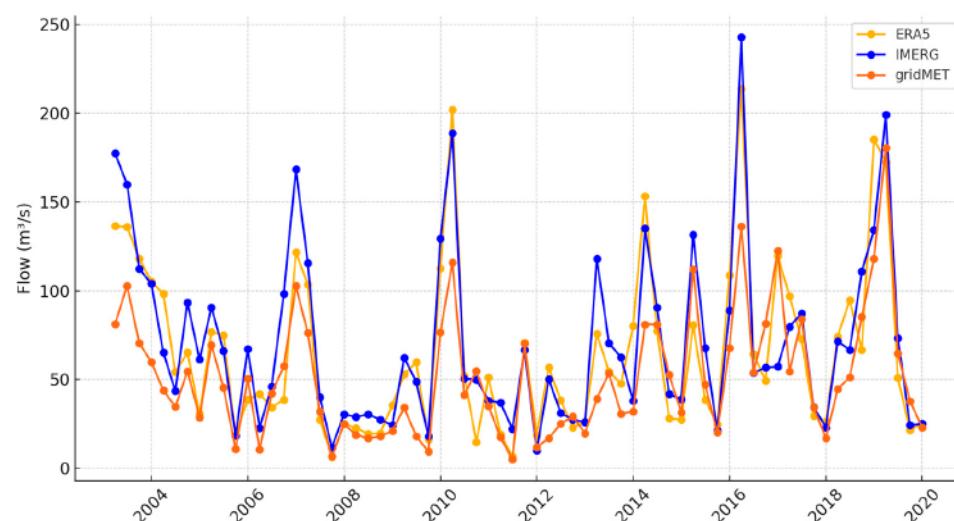

Figure 6. Seasonal average flow values at the outlet of Lower Tar sub-watershed for three rainfall datasets [The plot illustrates the flow patterns for ERA5, IMERG, and gridMET rainfall datasets, showing distinct seasonal peaks with variability across the years].

In summary, the seasonal flow analysis based on these rainfall datasets shows a strong relationship between seasonal rainfall patterns and flow dynamics in Washington, NC. Each dataset provides unique insights into the hydrological responses to different rainfall models, with IMERG consistently capturing higher rainfall-driven flows, while gridMET presents a more moderated view. This seasonal variability highlights the importance of considering multiple rainfall datasets when analyzing water flow for planning and management purposes [10,30–35].

3.3. Nitrate Load Variation Across Locations and Seasons

The annual average nitrate load values (Figure 7) reveal a consistent pattern where IMERG generally predicts lower nitrate loads across all locations, while ERA5 and gridMET consistently estimate the highest nitrate loads across most locations and seasons[6]. Pamlico Sound exhibits the highest nitrate load across all datasets, with gridMET estimating approximately 3850 metric tons/year, followed by ERA5 at 3750 metric tons/year and IMERG at around 1750 metric tons/year. This trend continues in other regions, such as Pamlico, where gridMET and ERA5 estimate a load of around 3410 and 3400 metric tons/year, respectively, compared to IMERG's 1580 metric tons/year. These differences can be attributed to the fact that the SWAT+ model was calibrated using IMERG data [10], which may lead to more conservative estimates of nitrate transport compared to other datasets, particularly in regions with more variable or extreme rainfall patterns [30][7].
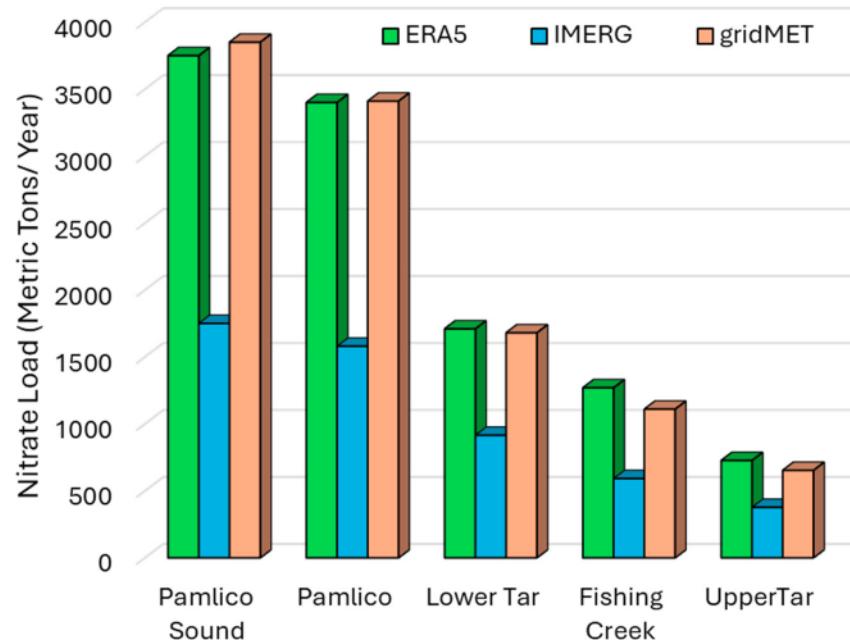

Figure 7. Annual average nitrate load (metric tons/year) across five sub-watersheds in the Tar-Pamlico Basin, comparing predictions from three rainfall datasets: ERA5, IMERG, and gridMET [gridMET consistently estimates the highest nitrate loads, while IMERG predicts significantly lower values across all locations].

In the case of the Lower Tar, ERA5 and gridMET produce similar trends, with gridMET being slightly lower at 1680 metric tons/year and ERA5 at 1710 metric tons/year, while IMERG predicts significantly lower values at 915 metric tons/year. This pattern demonstrates the dataset variability and the significant role of rainfall inputs in influencing nitrate transport estimates. The lower nitrate load predicted by IMERG can be attributed to the calibration of the SWAT+ model using IMERG rainfall data [10], which tends to represent lower or more diffuse rainfall patterns. Although IMERG produces higher flow predictions, its lower nitrate load may be due to the calibration of nitrate-related parameters in the SWAT+ model, which was optimized for IMERG's rainfall characteristics. This calibration

might have altered the model's sensitivity to nutrient transport processes, leading to reduced nitrate loads despite higher flows [10]. These differences highlight the importance of selecting appropriate rainfall datasets for hydrological modeling, as the choice can substantially affect nitrate load predictions [9,29–31,36–43].

The annual average nitrate load maps (Figure 8) for ERA5, gridMET, and IMERG reveal distinct patterns in nitrate transport across the Tar-Pamlico watershed, with significant variability between the datasets. ERA5 and gridMET predict the highest nitrate loads in the downstream subbasins, particularly in the Pamlico and Pamlico Sound regions[8], where the nitrate loads exceed 3000 metric tons per year. GridMET shows a more extensive area in the highest nitrate load class (>3000 metric tons/year), while ERA5 also shows large downstream regions with high nitrate transport. IMERG, on the other hand, predicts significantly lower nitrate loads across most of the watershed, with only limited areas in the Pamlico Sound exceeding 1700 metric tons/year[9].
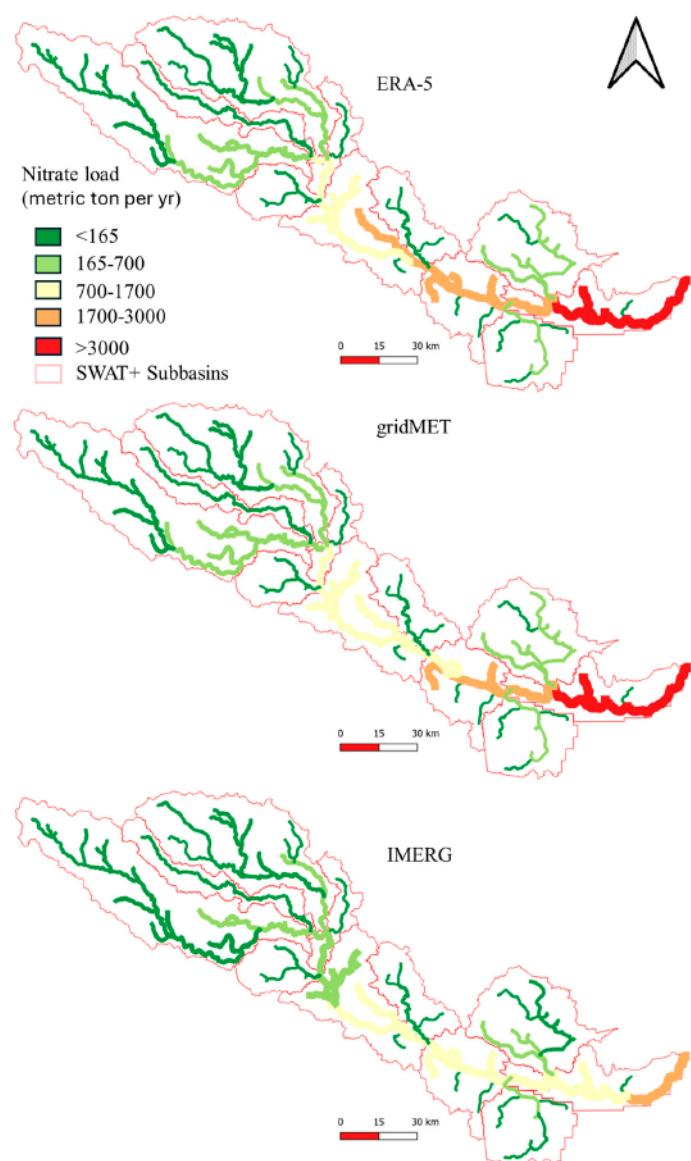
Figure 8. Spatial distribution of annual nitrate load across SWAT+ delineated channels using ERA5, gridMET, and IMERG datasets [The maps show nitrate load (metric tons per year) categorized into five classes, highlighting variability in nitrate transport estimates among the datasets across the watershed].

This study analyzed the variation in the average seasonal nitrate load across five regions in the Tar-Pamlico watershed—Upper Tar, Fishing Creek, Lower Tar, Pamlico, and Pamlico Sound—using IMERG rainfall datasets, and the model was calibrated to capture nitrate transport dynamics effectively across varying seasonal conditions using IMERG dataset [10]. The seasonal analysis of nitrate load across regions reveals clear temporal patterns, driven largely by the interaction of rainfall intensity, flow dynamics, and agricultural runoff. Winter and spring consistently exhibit the highest nitrate loads across all locations, while summer and fall generally see a reduction in monthly nitrate loads. These seasonal trends are crucial for understanding how nutrient runoff varies throughout the year and for identifying critical periods for water management policies.

During the winter months (December to February), despite lower rainfall, nitrate loads are relatively high across the watershed, particularly in downstream areas like Pamlico Sound. This season's reduced evapotranspiration allows more water to remain in the system, leading to higher flows and greater nitrate transport. For example, Pamlico Sound experiences an average nitrate load of 347,000 kg NO_3-N during winter. The high flow volumes during this period flush nitrates from agricultural fields and urban landscapes into the river system, resulting in elevated nutrient loads downstream [10,15].

Spring (March to May) also has higher nitrate transport, coinciding with increased rainfall and agricultural activities. The combination of spring rains and fertilizer application contributes to significant nitrate runoff, with downstream regions like Pamlico Sound showing the highest average nitrate load of 404,000 kg NO_3-N. This is a critical time for nutrient management, as the high runoff rates can exacerbate pollution risks, particularly in sensitive aquatic environments. Upstream regions, such as Fishing Creek and Lower Tar, also exhibit high nitrate loads during spring, reflecting the cumulative nutrient transport through the watershed.

In contrast, summer (June to August) sees a marked decrease in nitrate loads across all locations, despite high rainfall. This reduction can be attributed to higher evapotranspiration rates, which limit the amount of water available for runoff. Nitrate transport diminishes significantly, with Pamlico Sound experiencing an average load of 244,000 kg NO_3-N, down from the spring peak. However, localized nitrate concentrations may increase due to reduced dilution, necessitating careful monitoring of water quality during the summer months[10].

Fall (September to November) brings a moderate recovery in nitrate loads as flow increases, particularly in downstream regions. Pamlico and Pamlico Sound, for example, averaged 301,000 kg NO_3-N and 320,000 kg NO_3-N, respectively, during this period. Fall represents a transition period where rainfall helps recharge the watershed, preparing for the higher flow volumes of the upcoming winter months.

Figure 9 represents the time series of monthly nitrate load in Washington, NC, for the three rainfall datasets (ERA5, IMERG, and gridMET), illustrating clear trends and variations across the datasets, with IMERG in dark blue, indicating a tendency to predict lower nitrate loads across all seasons. Meanwhile, ERA5 and gridMET follow different patterns, showing noticeable variability across seasons.

All three datasets—ERA5, IMERG, and gridMET—display significant variability in nitrate loads across the seasons. Typically, nitrate loads peak during winter and spring, coinciding with periods of lower evapotranspiration and higher runoff. In contrast, lower nitrate loads are observed during the summer and fall, even with increased agricultural activities. This implies that environmental factors, such as runoff and evapotranspiration, have a greater influence on nitrate loads than agricultural practices alone. Throughout most observed periods, IMERG consistently shows lower nitrate loads, with ERA5 showing higher nitrate loads. This pattern indicates that the ERA5 dataset may be more responsive to heavy rainfall events, which drive greater nitrate runoff.
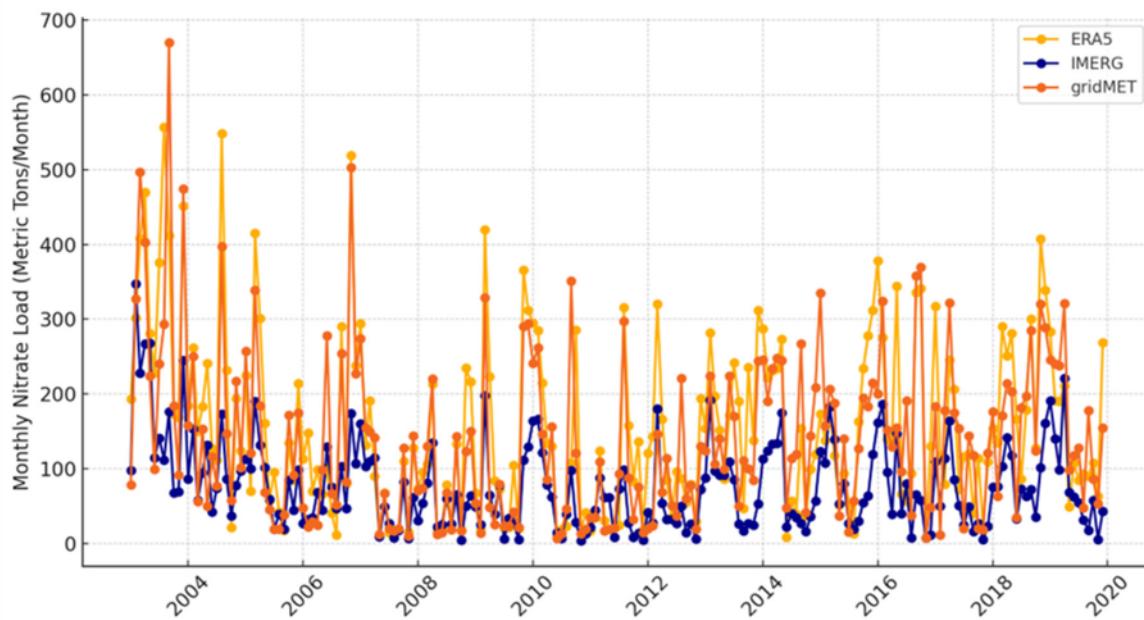
Figure 9. Comparison of monthly nitrate load (metric tons/month) from three datasets: ERA5, IMERG, and gridMET from 2003 to 2019 [The data show significant variations, with peak loads occurring in different periods for each dataset].

3.4. Flow–Nitrate Relationship

In this study, the flow–nitrate relationship reveals some surprising dynamics across the Tar-Pamlico watershed, driven by the use of different rainfall datasets—IMERG, ERA5, and gridMET. IMERG predicts higher rainfall and runoff and surprisingly estimates lower nitrate loads compared to ERA5 and gridMET[11]. This counterintuitive result can be attributed to the calibration of the SWAT+ model for the IMERG dataset, where the nitrate-related parameters were adjusted to align with IMERG's rainfall characteristics, potentially leading to reduced sensitivity in nitrate transport predictions [10,15]. Seasonal variations also play a crucial role, with winter showing the highest runoff and nitrate loads due to low evapotranspiration, despite having less rainfall than summer. In contrast, summer experiences higher rainfall but lower runoff and nitrate transport, as much of the water is lost to evapotranspiration. ERA5, despite predicting lower rainfall, estimates higher nitrate loads, suggesting it may capture more intense nitrate runoff events. GridMET, which predicts the lowest runoff, still shows substantial nitrate loads, particularly in winter.

These findings emphasize the complex and sometimes surprising interactions between flow and nitrate transport, where dataset selection and model calibration can significantly affect predictions, especially during critical periods such as winter (high runoff and nitrate load) and fall (low nitrate load). This variability underscores the importance of considering multiple rainfall datasets when evaluating nitrate loads for nutrient management and water quality planning [10,31,44–52].

4. Limitations and Future Research Directions

A key limitation of this study is that the SWAT+ model was only calibrated and validated using the IMERG rainfall dataset, which may introduce bias in comparisons with ERA5 and gridMET. Since the model is optimized for IMERG's characteristics, it could lead to underestimation or overestimation of flow and nitrate transport when using other datasets. Future research should focus on multi-dataset calibration and validation to reduce these biases and provide a more accurate comparison. Additionally, using finer temporal scales (e.g., daily or hourly) and incorporating local gauge or high-resolution satellite data could further improve hydrological predictions.

5. Conclusions

This study has provided a thorough comparative analysis of three rainfall datasets—ERA5, IMERG, and gridMET—in simulating monthly river flow and monthly nitrate load within the Tar-Pamlico watershed using the SWAT+ model. The results reveal significant variability in both flow and nitrate load predictions, demonstrating the complex interactions between rainfall, runoff, and nutrient transport in a coastal agricultural watershed. IMERG consistently predicts higher rainfall and runoff, while gridMET and ERA5 present more conservative rainfall estimates but higher nitrate loads, especially in downstream regions like Pamlico Sound. These variations underscore the importance of selecting appropriate rainfall datasets for hydrological modeling, as each dataset characteristic can significantly influence predictions of nutrient transport and environmental impacts.

Seasonal dynamics play a critical role in nitrate transport, with winter showing higher runoff and nitrate loads due to lower evapotranspiration and increased water retention in the system, despite receiving less rainfall compared to summer. Conversely, summer exhibits higher rainfall but lower runoff and nitrate load due to higher evapotranspiration rates, which limit water flow and nutrient movement. Seasonal dynamics suggest that flow drives nitrate load more than agricultural activities. Despite lower farming activity in winter, higher runoff and water retention lead to increased nitrate loads, indicating that water movement is the primary factor in transporting nitrates load rather than agricultural inputs alone.

The study's findings emphasize the need for environmental modelers to use multiple rainfall datasets to improve the robustness of hydrological models, particularly in vulnerable regions prone to eutrophication and hypoxia. The variability in predictions suggests that relying on a single dataset could lead to inaccurate assessments of nutrient pollution risks, potentially undermining mitigation efforts.

Notes

1. This study used satellite rainfall data from Tapas et al., 2023 [9], which examined the effects of different autocalibration techniques and three rainfall datasets—ERA5, IMERG, and gridMET—on daily SWAT+ flow simulations. Building on that work, this study further evaluates the impact of these rainfall datasets on riverine monthly nitrate load simulations.

2. SWAT+ delineated subbasins by Tapas et al., 2024a [10] and Tapas, 2024b [15]—which were used in this study—and HUC-8 boundaries did not properly align in this study. To address this, the downstream SWAT+ channel overlapping with each

3   corresponding HUC-8 boundary was used. Specifically, for the Pamlico Sound HUC-8, the SWAT+ model's final channel was located significantly upstream of the actual HUC-8 outlet for Pamlico Sound (Figure S1, Supplementary Information).

3   SWAT+ is a one-dimensional model, which means it does not account for backflows or complex hydrodynamics in downstream regions, such as the Pamlico and Pamlico Sound [10]. As a result, the model may overestimate flow values in these areas, where wider channels and tidal influences can significantly impact water movement. Additionally, the lack of observed data in the most downstream sections limits the ability to properly calibrate and validate the model's predictions for these regions, further contributing to potential discrepancies between the simulated and actual flow values [10].

4   The most downstream channel in SWAT+ is located considerably upstream of the Pamlico Sound's HUC-8 outlet point. This may result in relatively similar values being predicted for both the Pamlico and Pamlico Sound in this study.

5   SWAT+ is a one-dimensional model, which means it does not account for backflows or complex hydrodynamics in downstream regions, such as the Pamlico and Pamlico Sound [10]. As a result, the model may overestimate flow values in these areas, where wider channels and tidal influences can significantly impact water movement. Additionally, the lack of observed data in the most downstream sections limits the ability to properly calibrate and validate the model's predictions for these regions, further contributing to potential discrepancies between the simulated and actual flow values [10].

6   The SWAT+ model developed by Tapas et al. (2024a [10] and 2024b [15]) was originally calibrated using IMERG rainfall data. In this study, the same SWAT+ model, along with the calibrated parameters for IMERG data, was applied to evaluate nitrate loads using both gridMET and ERA5 datasets. However, the results could change significantly if the model was recalibrated specifically for the ERA5 and gridMET datasets, which is beyond the scope of this study.

7   The SWAT+ model developed by Tapas et al. (2024a [10] and 2024b [15]) was originally calibrated using IMERG rainfall data. In this study, the same SWAT+ model, along with the calibrated parameters for IMERG data, was applied to evaluate nitrate loads using both gridMET and ERA5 datasets. However, the results could change significantly if the model was recalibrated specifically for the ERA5 and gridMET datasets, which is beyond the scope of this study.

8   The most downstream channel in SWAT+ is located considerably upstream of the Pamlico Sound's HUC-8 outlet point. This may result in relatively similar values being predicted for both the Pamlico and Pamlico Sound in this study.

9   The SWAT+ model developed by Tapas et al. (2024a [10] and 2024b [15]) was originally calibrated using IMERG rainfall data. In this study, the same SWAT+ model, along with the calibrated parameters for IMERG data, was applied to evaluate nitrate loads using both gridMET and ERA5 datasets. However, the results could change significantly if the model was recalibrated specifically for the ERA5 and gridMET datasets, which is beyond the scope of this study.

10   It is important to note that this observation pertains specifically to nitrate load. Nitrate concentrations may exhibit significantly different patterns, as they can increase with reduced water flow. This study focuses on nitrate load as the model was calibrated for monthly nitrate load.

11   It is important to note that these values are specific to nitrate load, as the model was calibrated for this variable [10,15], and nitrate concentration could be entirely different. In fact, nitrate concentrations might be higher in fall and summer due to lower flows and ongoing agricultural activities, which can result in less dilution of the nitrates present in the water.

References

1. Montagna, P.; Palmer, T.A.; Pollack, J.B. *Hydrological Changes and Estuarine Dynamics*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
2. Tapas, M.R.; Kumar, U.; Mogili, S.; Jayakumar, K.V. Development of multivariate integrated drought monitoring index (MIDMI) for Warangal region of Telangana, India. *J. Water Clim. Chang.* **2022**, *13*, 1612–1630. [CrossRef]
3. Mankar, T.S.; Mane, S.; Mali, S.T.; Tapas, M.R. Analysis and Development of Watershed for Ruikhed Village, Maharashtra—A Case Study. *Int. Res. J. Eng. Technol.* **2020**, *7*, 2265–2270.
4. Sharma, N.N.; Tapas, M.R.; Kumar, A.U. Drought Monitoring Indices. *Meteorol. Climatol.* **2022**, *53*, 53–59.
5. Myers, D.T.; Ficklin, D.L.; Robeson, S.M. Incorporating rain-on-snow into the SWAT model results in more accurate simulations of hydrologic extremes. *J. Hydrol.* **2021**, *603*, 126972. [CrossRef]
6. Prabha, J.A.; Tapas, M.R. Event-based rainfall-runoff modeling using HEC-HMS. *IOSR J. Mech. Civ. Eng.* **2020**, *17*, 41–59.
7. Mishra, G.J.; Kumar, A.U.; Tapas, M.R.; Oggu, P.; Jayakumar, K.V. Evaluating hydrological alterations and recommending minimum flow release from the Ujjani dam to improve the Bhima River ecosystem health. *Water Sci. Technol.* **2023**, *88*, 763–777. [CrossRef]
8. Pesce, M.; Critto, A.; Torresan, S.; Giubilato, E.; Santini, M.; Zirino, A.; Ouyang, W.; Marcomini, A. Modelling climate change impacts on nutrients and primary production in coastal waters. *Sci. Total Environ.* **2018**, *628*, 919–937. [CrossRef]
9. Tapas Mahesh, R.; Etheridge, R.; Tran, T.-N.-D.; Hung, M.; Hinckley, B.; Nguyen, T.; Lakshmi, V. Satellite-based rainfall datasets and autocalibration techniques' effects on SWAT+ flow prediction. *Authorea* **2023**. [CrossRef]
10. Tapas Mahesh, R.; Etheridge, R.; Colin, G.F.; Ariane, L.P.; Bell, N.; Xu, Y.; Lakshmi, V. A methodological framework for assessing sea level rise impacts on nitrate loading in coastal agricultural watersheds using SWAT+: A case study of the Tar-Pamlico River basin, North Carolina, USA. *Sci. Total Environ.* **2024**, *951*, 175523. [CrossRef]

11. Tapas, M.; Etheridge, J.R.; Howard, G.; Lakshmi, V.V.; Tran, T.N.D. Development of a Socio-Hydrological Model for a Coastal Watershed: Using Stakeholders' Perceptions. In Proceedings of the AGU Fall Meeting Abstracts, San Francisco, CA, USA, 11–16 December 2022; Volume 2022, p. H22O-0996.

12. Tran, T.N.D.; Tapas, M.; Tran, V.; Arshad, A.; Nguyen, B.Q.; Do, S.K.; Etheridge, J.R. Comparison of SWAT and SWAT+: Review and Recommendations. In Proceedings of the AGU Fall Meeting Abstracts, San Francisco, CA, USA, 11–15 December 2023; AGU23.

13. Yin, D.; Harris, C.; Tran, T.N.D.; Tapas, M.; Etheridge, J.R.; Moysey, S.M.; Lakshmi, V.V. Effects of Sea-Level Rise and River Flow Variation on Estuarine Salinity in a Changing Climate: Insights from the Pamlico River Estuary, USA. In Proceedings of the 2024 Ocean Sciences Meeting, New Orleans, LA, USA, 18–23 February 2024; AGU.

14. Tran, T.-N.-D.; Mahesh, R.T.; Son, K.D.; Etheridge, R.; Lakshmi, V. Investigating the impacts of climate change on hydroclimatic extremes in the Tar-Pamlico River basin, North Carolina. *J. Environ. Manag.* **2024**, *363*, 121375. [CrossRef]

15. Tapas, M. Integrative Analysis of Policy Changes for a Coastal Watershed: Implications for Agriculture and Ecosystem Health. Ph.D. Thesis, East Carolina University, Greenville, NC, USA, 2024. Available online: http://hdl.handle.net/10342/13413 (accessed on 1 September 2024).

16. Luchette, J.A.; Crawford, T. A public participation GIS application for citizen-based watershed monitoring in the Pamlico-Tar River basin, North Carolina. *Southeast. Geogr.* **2008**, *48*, 184–200. [CrossRef]

17. Heffernan, J. Spatial and Temporal Analysis of Long-Term Water Quality Data for the Pamlico River Estuary, North Carolina. Ph.D. Thesis, Duke University, Durham, NC, USA, 2015.

18. McAllister, D.E.; Hamilton, A.L.; Harvey, B. Global freshwater biodiversity: Striving for the integrity of freshwater ecosystems. *Sea Wind. Bull. Ocean. Voice Int.* **1997**, *11*.

19. Costa, D.; Sutter, C.; Shepherd, A.; Jarvie, H.; Wilson, H.; Elliott, J.; Liu, J.; Macrae, M. Impact of climate change on catchment nutrient dynamics: Insights from around the world. *Environ. Rev.* **2022**, *31*, 4–25. [CrossRef]

20. Johnston, R.; Smakhtin, V. Hydrological modeling of large river basins: How much is enough? *Water Resour. Manag.* **2014**, *28*, 2695–2730. [CrossRef]

21. García, L.; Rodríguez, D.; Wijnen, M.; Pakulski, I. (Eds.) *Earth Observation for Water Resources Management: Current Use and Future Opportunities for The Water Sector*; World Bank Publications: Chicago, IL, USA, 2016.

22. Pradhan, R.K.; Markonis, Y.; Godoy, M.R.; Villalba-Pradas, A.; Andreadis, K.M.; Nikolopoulos, E.I.; Papalexiou, S.M.; Rahim, A.; Tapiador, F.J.; Hanel, M. Review of GPM IMERG performance: A global perspective. *Remote Sens. Environ.* **2022**, *268*, 112754. [CrossRef]

23. Wolkeba, F.T.; Mekonnen, M.M. Evaluation of gridded precipitation data in water availability modeling in CONUS. *J. Hydrol.* **2024**, *628*, 130575. [CrossRef]

24. Lin, Z.; Yao, X.; Du, J.; Zhou, Z. Refined evaluation of satellite precipitation products against rain gauge observations along the Sichuan—Tibet railway. *J. Meteorol. Res.* **2022**, *36*, 779–797. [CrossRef]

25. Giordani, A.; Cerenzia, I.M.; Paccagnella, T.; Di Sabatino, S. SPHERA, a new convection-permitting regional reanalysis over Italy: Improving the description of heavy rainfall. *Q. J. R. Meteorol. Soc.* **2023**, *149*, 781–808. [CrossRef]

26. NCDEQ (North Carolina Department of Environmental Quality). *Tar-Pamlico River Basin 2010 Water Quality Plan*; NCDEQ: Raleigh, NC, USA, 2010.

27. Phillips, J.D. River discharge and sediment deposition in the Upper Pamlico Estuary. In *Estuarine Circulation*; Humana Press: Totowa, NJ, USA, 1989; pp. 337–349.

28. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. [CrossRef]

29. Sayemuzzaman, M.; Jha, M.K. Seasonal and annual precipitation time series trend analysis in North Carolina, United States. *Atmos. Res.* **2014**, *137*, 183–194. [CrossRef]

30. Boyles, R.P.; Raman, S. Analysis of climate trends in North Carolina (1949–1998). *Environ. Int.* **2003**, *29*, 263–275. [CrossRef] [PubMed]

31. Álvarez, X.; Valero, E.; Santos, R.M.; Varandas, S.G.; Fernandes, L.S.; Pacheco, F.A. Anthropogenic nutrients and eutrophication in multiple land use watersheds: Best management practices and policies for the protection of water resources. *Land Use Policy* **2017**, *69*, 1–11. [CrossRef]

32. Tran, T.N.D.; Nguyen, Q.B.; Vo, N.D.; Marshall, R.; Gourbesville, P. Assessment of Terrain Scenario Impacts on Hydrological Simulation with SWAT Model. In *Application to Lai Giang Catchment, Vietnam*; Springer Water, Springer Nature: Berlin/Heidelberg, Germany, 2022; pp. 1205–1222. [CrossRef]

33. Ahmed, Z.; Tran, T.N.D.; Nguyen, Q.B. Applying Semi Distribution Hydrological Model SWAT to Assess Hydrological Regime in Lai Giang Catchment, Binh Dinh Province, Vietnam. In *Proceedings of the 2nd Conference on Sustainability in Civil Engineering (CSCE'20, August 2020)*; Capital University of Science and Technology: Islamabad, Pakistan, 2020; p. 8. Available online: https://csce.cust.edu.pk/archive/20-404.pdf (accessed on 9 July 2024).

34. Nguyen, B.Q.; Vo, N.D.; Le, M.H.; Nguyen, Q.D.; Lakshmi, V.; Bolten, J.D. Quantification of Global Digital Elevation Model (DEM)—A Case Study of the Newly Released NASADEM for a River Basin in Central Vietnam. *J. Hydrol. Reg. Stud.* **2023**, *45*, 101282. [CrossRef]

35. Do, S.K.; Nguyen, B.Q.; Tran, V.N.; Grodzka-Lukaszewska, M.; Sinicyn, G.; Lakshmi, V. Investigating the Future Flood and Drought Shifts in the Transboundary Srepok River Basin Using CMIP6 Projections. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 7516–7529. [CrossRef]

36. Tran, T.-N.-D.; Nguyen, B.Q.; Zhang, R.; Aryal, A.; Grodzka-Łukaszewska, M.; Sinicyn, G.; Lakshmi, V. Quantification of Gridded Precipitation Products for the Streamflow Simulation on the Mekong River Basin Using Rainfall Assessment Framework: A Case Study for the Srepok River Subbasin, Central Highland Vietnam. *Remote Sens.* **2023**, *15*, 1030. [CrossRef]

37. Tran, T.-N.-D.; Nguyen, B.Q.; Grodzka-Łukaszewska, M.; Sinicyn, G.; Lakshmi, V. The Role of Reservoirs under the Impacts of Climate Change on the Srepok River Basin, Central Highlands of Vietnam. *Front. Environ. Sci.* **2023**, *11*, 1304845. [CrossRef]

38. T Le, M.-H.; Zhang, R.; Nguyen, B.Q.; Bolten, J.D.; Lakshmi, V. Robustness of Gridded Precipitation Products for Vietnam Basins Using the Comprehensive Assessment Framework of Rainfall. *Atmos. Res.* **2023**, *293*, 106923. [CrossRef]

39. Aryal, A.; Tran, T.; Kumar, B.; Lakshmi, V. Evaluation of Satellite-Derived Precipitation Products for Streamflow Simulation of a Mountainous Himalayan Watershed: A Study of Myagdi Khola in Kali Gandaki. *Remote Sens.* **2023**, *15*, 4762. [CrossRef]

40. Lakshmi, V. Enhancing Human Resilience against Climate Change: Assessment of Hydroclimatic Extremes and Sea Level Rise Impacts on the Eastern Shore of Virginia, United States. *Sci. Total Environ.* **2024**, *947*, 174289. [CrossRef]

41. Lakshmi, V. Visualization-driven Hydrologic Assessment Using Gridded Precipitation Products. *Hydrol. Process.* **2024**, *38*, e15286. [CrossRef]

42. Nguyen, B.Q.; Kantoush, S.A.; Tran, T.-N.-D.; van Binh, D.; Vo, N.D.; Saber, M.; Sumi, T. Response of Hydrological to Anthropogenic Activities in a Tropical Basin. In Proceedings of the 40th IAHR World Congress, Vienna, Austria, 25 August 2023; pp. 269–278. [CrossRef]

43. Nguyen, B.Q.; Van Binh, D.; Tran, T.-N.-D.; Kantoush, S.A.; Sumi, T. Response of Streamflow and Sediment Variability to Cascade Dam Development and Climate Change in the Sai Gon Dong Nai River Basin. *Clim. Dyn.* **2024**, *62*, 7997–8017. [CrossRef]

44. Tamagno, S.; Eagle, A.J.; McLellan, E.L.; van Kessel, C.; Linquist, B.A.; Ladha, J.K.; Lundy, M.E.; Pittelkow, C.M. Predicting nitrate leaching loss in temperate rainfed cereal crops: Relative importance of management and environmental drivers. *Environ. Res. Lett.* **2022**, *17*, 064043. [CrossRef]

45. Alex, N.; Dong, F.; Shimoda, Y.; Arnillas, C.A.; Javed, A.; Yang, C.; Zamaria, S.; Mandal, S.; Wellen, C.; Paredes, D.; et al. A review of the current state of process-based and data-driven modelling: Guidelines for Lake Erie managers and watershed modellers. *Environ. Rev.* **2021**, *29*, 443–490.

46. Ghannem, S.; Paredes-Arquiola, J.; Bergillos, R.J.; Solera, A.; Andreu, J. Assessing the Effects of Environmental Flows on Water Quality for Urban Supply. *Water* **2024**, *16*, 1509. [CrossRef]

47. Amos, H.M.; Miniat, C.F.; Lynch, J.; Compton, J.; Templer, P.H.; Sprague, L.A.; Shaw, D.; Burns, D.; Rea, A.; Whitall, D.; et al. What goes up must come down: Integrating air and water quality monitoring for nutrients. *Environ. Sci. Technol.* **2018**, *52*, 11441–11448. [CrossRef]

48. Behrangi, A.; Khakbaz, B.; Jaw, T.C.; AghaKouchak, A.; Hsu, K.; Sorooshian, S. Hydrologic evaluation of satellite precipitation products over a mid-size basin. *J. Hydrol.* **2011**, *397*, 225–237. [CrossRef]

49. Tang, G.; Clark, M.P.; Papalexiou, S.M.; Ma, Z.; Hong, Y. Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. *Remote Sens. Environ.* **2020**, *240*, 111697. [CrossRef]

50. Hirpa, F.A.; Gebremichael, M.; Hopson, T. Evaluation of high-resolution satellite precipitation products over very complex terrain in Ethiopia. *J. Appl. Meteorol. Climatol.* **2010**, *49*, 1044–1051. [CrossRef]

51. Liu, C.Y.; Aryastana, P.; Liu, G.R.; Huang, W.R. Assessment of satellite precipitation product estimates over Bali Island. *Atmos. Res.* **2020**, *244*, 105032. [CrossRef]

52. Todeschini, S. Innovative and Reliable Assessment of Polluted Stormwater Runoff for Effective Stormwater Management. *Water* **2024**, *16*, 16. [CrossRef]