Summary of Responses to the Editor

1. All the comments from both reviewers have been addressed/responded in detail.
2. Percentage area under agriculture and urban land use have been computed for each watershed based on the NLCD data of 2001. This clarifies the point some of the watersheds have significant agricultural activity within the basin.
3. A new figure (Figure 9) has been added to quantify the role of each predictor, forecasted precipitation and 3-day average streamflow, in influencing the forecast skill. This also provides information on the role of basin storage in forecasting the TN loadings.
4. Another figure (Figure 10) has been added to show the role of land use type particularly the area under agriculture in influencing the skill of the forecast. This analyses shows that watersheds with higher percentage of agricultural land exhibit higher skill. This indicates the role of basin nutrient storage in influencing the forecast skill.
5. Detailed responses/corrections have been provided to the suggestions by reviewer-2.

Thanks for handling the manuscript.

Reviewer # 1

1. ***Comment:*** *Generally, the paper is well-written and the methodology is clearly presented. However, there are major concerns with regard to the contribution of the paper: Firstly, application of the model is quite limited. Nutrient loadings have been a major concern in agricultural and environmental engineering because of human use of fertilizers. On the other hand, for undeveloped basins, water quality is usually quite good and nutrient loading is not a problem. Given that the model can only be applied to undeveloped basins, it is concluded that the paper tries to but actually fails to address the important issue of nutrient loadings. To make it a solid contribution, the paper should develop a model for river basins subject to human interferences.*

   Response: Thanks for this comment. This is a good point and requires clarification. We should have written it more clearly. HCDN stations are undeveloped watersheds from hydrology perspective with streamflow from the station being not influenced by upstream storage or groundwater pumping. This is needed from the perspective linking weather to streamflow. Otherwise, the flows will be controlled due to operational guidelines of the reservoir and we will not be relate it to the weather information.

   To substantiate this point, we have added the percentage area under agriculture in Table 1 based on the National Land Use Classification Data (NLCD) data of 2001. From Table 1, we can see the distribution with seven, six and five watersheds having 20%-30%, 10%-20% and 0%-10% of area under agriculture respectively. Thus, the watersheds are not completely undeveloped without any agricultural activity.

   We have revised the manuscript based on the above response.

2. ***Comment:*** *Secondly, the simple combination of two models doesn't present a novel contribution. The model for nutrient loading forecast is based on the k-NN model and the LOADEST model, which are classical models and have been applied to many cases. To make the model combination a solid contribution, the paper should exploit the models and derive some new understandings, e.g., structural relationships between daily streamflow and nutrient loading.*

   Response: To address this, we have added Figure 9 that quantifies the role of different predictors, 3-day average streamflow prior to forecasting day and 1-day ahead precipitation forecasts, on the overall skill in forecasting TN loadings for all the 18 sites in the Southeast. Figure 9 clearly indicates that the combination of both 3-day average streamflow and 1-day ahead precipitation forecasts as predictors result in improved correlation and reduced RMSE in estimating daily TN loadings at all the sites. Comparing the skill obtained from 3-day average streamflow and forecasted precipitation, we infer that for most of the watersheds, the skill obtained using 3-day average streamflow (prior to the forecasting day) alone as a predictor provides better skill in comparison to the skill obtained using forecasted precipitation alone as a predictor. But, including both predictors result in overall improvement. We have included details related to this under discussion.

3. ***Comment:*** *Thirdly, besides daily streamflow, nutrient storage in the river basin is another key determinant of nutrient loading. This paper didn't consider this issue in nutrient loading forecast. Notably, the statistical models k-NN and LOADEST are based on historical samples, and the underlying assumption is "stationarity". However, nutrient storage can vary with time and is greatly affected by human interferences. How to consider this kind of non-stationarity in the statistical model?*

   Response: This comment has two parts: First, the issue of nutrient storage in the river reach. Though instream loadings primarily depend on streamflow and precipitation variability during the season, antecedent moisture/flow conditions also play a critical role in influencing the nutrient loadings from the watershed (Vecchia 2003, Alexander and Smith 2006). Our hypothesis is that by considering 3-day average streamflow prior to the forecasting day accounts for the storage in the river. Increased streamflow indicates increased streamflow storage conditions in the basin, which could potentially increase nutrient storage in the river. This is further validated by our plot added (Figure 9). From Figure 9, storage in the river reach explain higher variability in the observed nutrients as opposed to the skill obtained using the forecasted precipitation alone.

   With regard to the second question on "non-stationarity", Oh and Sankarasubramanian (2012) clearly show that there is no trend in the observed TN nutrients in the WQN data over the 18 stations. Given that the basins are from the HCDN basins, the streamflow is also not influenced by upstream storage or pumping with no exhibited time trend. Thus, the LOADEST model is appropriate for the time period that we have considered. Thus, the presented modeling framework is applicable for the selected stations and for the evaluation period.

   However, developing a model for basin experiencing significant land use changes due to human influences require additional information. For instance, if the basin experiences significant urbanization, then it is natural to expect the point TN loadings from waste water treatment (WWT) plants to influence the downstream loadings and concentration. Under such situation, we need to know the discharges from the WWT plants to develop the model.

   On the other hand, for a basin experiencing significant changes in agricultural land use, it may be important to know the area under cultivation for each year so that it could be considered as an additional predictor. To substantiate this point, we investigated relating (Figure 10) the nutrient forecast ability to the percentage of area under agriculture provided in Table 1. Based on that, we infer that as the percentage area under agriculture increases the skill in forecasting nutrients. This indicates that as the nutrient storage in the basin increases, the transport induced by the streamflow increases resulting in improved skill in forecasting the nutrients. Any non-stationarity in the predictand could be explained if the appropriate time-varying predictor is identified. Information from remote sensing satellites that quantify the Chlorophyll concentration could be also be considered as nutrient storage in the river reach and water bodies.

To our knowledge, this is the first study that evaluates the ability to forecast nutrients based on actual streamflow forecasts. It is beyond the scope of this study to consider basins experiencing significant human interferences. However, we have added the above points in the discussion for forecasting TN for basins experiencing significant human interferences.

Thanks for the detailed comments!

Reviewer # 2

1. ***Comment:*** *My In "Introduction", the reasons why the streamflow over the previous three days could be considered as a surrogate for antecedent soil moisture conditions need to be further clarified in the revised manuscript.*
   Response: We have added few references to support that antecedent streamflow conditions can be considered as surrogates for initial conditions of the catchments. We added in lines 93-94: "Furthermore, several studies have utilized antecedent streamflows as surrogates of initial catchment conditions (e.g. Chiew and McMohan, 2002; Piechota et al., 2001; Wang et al., 2009) including 3-day average streamflow conditions (Majumdar and Kumar, 1990; Srinivas and Srinivasan, 2000; Cigizoglu, 2003)." In our analysis, we found that the maximum correlation between observed streamflow and the previous-day streamflow seems to occur over 3 days for the selected sites.

2. ***Comment:*** *The 2th paragraph of Introduction, sentence "Developing daily streamflow forecasts over a large region using semi-distributed models require intensive spatial data (e.g. topography, land cover, soils) and computational resources, hence, we employed the KNearest Neighbor (K-NN) semi-parametric approach to develop daily streamflow forecasts contingent on updated climate forecasts." This sentence needs to be rewritten in a clearer way to highlight the reason that K-NN resampling approach was applied in this study.*
   Response: We have updated the text in lines 84-91 to clarify this as: "Developing daily streamflow forecasts over a large region using semi-distributed models require intensive spatial data (e.g. topography, land cover, soils) and computational resources, hence, we employed a semi-parametric approach in this study. In particular, we employed the K Nearest Neighbor (K-NN) semi-parametric scheme to develop daily streamflow forecasts contingent on updated climate forecasts since it can capture nonlinear relationships that are typically observed in daily streamflow data (Salas and Lee, 2010). The K-NN scheme has been widely used in hydrologic studies (Rajagopalan and Lall 1999; Prairie et al. 2006; Sharif and Burn 2006)."

3. ***Comment:*** *In "2.2 Weather forecasts database": please present the mechanism of the method for forecasting precipitation.*
   Response: We added further details in lines 129-139 as: "The GFS forecasts model has 28 sigma (pressure) levels and a T62 spatial resolution (~200 km grid size), which represents physical processes to estimate atmospheric forcings such as winds, temperature, precipitation, geopotential heights at different pressure levels (Hamill et al., 2006). 15 ensemble forecasts are obtained by initializing different atmospheric states of the GFS model every day. The control run is initialized by the National Center for Environmental Prediction (NCEP)-National Center for Atmospheric Research (NCAR) reanalysis data (Kalnay et al., 1996) while the other 14 ensemble members use a set of seven bred pairs of initial conditions (Toth and Kalnay, 1997), which are re-centered each day on the reanalysis initial condition.

In this study, we make use of daily precipitation reforecasts from the GFS model consists of 15 ensemble members, up to 15 days in advance, starting from 1979 to till date."

4. *Comment: In "3. Stream and total nitrogen forecasting models", three models were proposed for forecasting precipitation, streamflow and nutrient loading, respectively. Do the model scales match? I suggest more descriptions on how these forecasting model scales match each other be presented.*

   Response: The time scales of the models match, but the spatial scales don't match. Hence, we considered a statistical model for developing streamflow forecast. We added the following sentence to address this concern: "The LODEST model can be employed with the observed or predicted daily streamflow time series at any given site. Streamflow forecast developed using large-scale precipitation forecasts and previous 3-day average streamflow using the non-parametric model is forced with the LOADEST model to develop nutrient forecasts."

5. *Comment: In "3.2 K-Nearest Neighbor (K-NN) resampling approach", there should be more description on the specific mechanism of K-NN resampling approach.*

   Response: Lines 194-223 describe the mechanism of the K-NN resampling approach. We revised the beginning to the paragraph to reflect this as: In the K-NN scheme, we used the Mahalanobis distance instead of the Euclidean distance, since the selected predictors – PCs of the principal components and the streamflow over the past observations- could be correlated."

6. *Comment: In "4.1 Skill in forecasting daily streamflow", I would suggest that the obtained conditional distribution of flows for 18 watersheds should be presented.*

   Response: It will be too many plots/data to show the conditional distribution of flows obtained for each forecasting day for all the 18 watersheds. However, we added Figure 9 to quantify the role of different predictors such as previous 3-day average streamflow prior to forecasting day and 1-day ahead precipitation forecasts in improving the skill in forecasting TN loadings for all the 18 sites. Figure 9 clearly indicates that the combination of both daily streamflow previous 3-days and 1-day ahead precipitation forecasts as predictors result in improved correlation and reduced RMSE in estimating daily TN loadings at all the sites.

7. *Comment: The contents of "5. Summary and conclusions" section should be enhanced according to the actual results. Please emphasize innovations and important conclusions.*

   Response: We have extended the discussion as well as the summary and conclusions to incorporate the results from the addition of Figures 9 and 10 along with substantiation of results from t

8. *Comment: Sections There are several typographical and grammatical errors that need to be corrected, for example: (a) Page 15628: "But, availability of data on total nitrogen is limited with concentration is typically measured on a non-continuous basis."*
Response: We have updated the sentence.

9. *Comment: (b) Page 15628: "Similarly, considerable progress has been made in developing daily streamflow forecasts using both statistical models that consider both parametric and semi-distributed models."*
Response: We revised the sentence as: "Similarly, considerable progress has been made in developing daily streamflow forecasts using statistical models, e.g. parametric models (Rajagopalan and Lall 1999; Anderson et al., 2002; Salas and Lee, 2010), and semi-distributed watershed models (e.g., Clark and Hay, 2004; Mcenery et al., 2005, Georgakakos et al., 2010)."

10. *Comment: (c) Page 15630: "The WQN database comprises of water quality data from USGS monitoring networks from both large watersheds (National Stream Quality Accounting Network, NASQAN) and minimally developed watersheds (Hydrologic Benchmark Network, HBN)."*
Response: The sentence has been corrected as: "The WQN database comprises of water quality data from the USGS monitoring networks for large watersheds (National Stream Quality Accounting Network, NASQAN) as well as watersheds that are minimally developed (Hydrologic Benchmark Network, HBN)."

11. *Comment: (d) Page 15635: "These errors primarily occur due to the inability of the model to predict high values, which resulted in very high residuals."*
Response: This has been corrected as: "These errors primarily occur due to the inability of the model to predict high values, as indicated by very high residuals."

Comment: (e) Page 15639:"It is important to note that all the skill reported in Figs. 3–6 consider the ability to predicting exactly for those days with WQN observations."
Response: This has been corrected as: "It is important to note that all the skill reported in Figures 3-6 consider the ability to predicting those days when the WQN observations are available.

Thanks for the detailed comments!