

POINT-BY-POINT RESPONSE TO THE REVIEWS

<Anonymous Referee #1>

Reviewer's Comment 1:

Model description in section 2.4 and Tables 1 and 2. It is not clear which parameters are fixed (from literature and/or measurements) and which are calibrated: in table 1 all listed (new) parameters are fixed?

Response 1:

This is a valid comment and we plan to modify Table 1 to clarify this point.

In the row "Reference or Rationale", we will specify if the parameters were calibrated or taken/adapted from published values.

Reviewer's Comment 2:

Also in Table 2 the word "variable" I think is not properly used: it seems a mix between state variables, inputs (or forcings) and parameters.

Response 2:

Agreed, we will modify Table 2 specifying if the variables are prognostic or diagnostic or inputs/forcings.

Reviewer's Comment 3:

The implementation should be better explained here: in the abstract is written the model calibration was done at Marietta and it was done a spatial validation using the rest of stations. Why it was not done a temporal validation? In my opinion, it can affect the temporal extrapolability/predictability of the model. Or not?

Response 3:

Agreed, we will modify page 5686, lines 3-7 to read:

We simulated with LM3-TAN stream dissolved organic-N, ammonium-N, and nitrate-N loads throughout the river network. The model was calibrated by comparing the modeled stream N loads with the corresponding reported N loads at the last downstream SRBC station Marietta, in which contributions of the entire watershed to the stream flows and N loads can be assessed. Thus, temporal evaluation of the stream discharges and N loads for the period 1987-2005 was focused on at the Marietta station. River data from the 15 monitoring stations (1986-2005) were also used to evaluate spatial stream discharges and N loads.

The temporal evaluation was done at Marietta. This is explained in the above response; page 5687, lines 3-5; Fig. 5.

Reviewer's Comment 4:

Section 7 is relatively short. I miss results concerning the implementation and exploitation of hydrological state variables. Probably there are interactions with N state variables.

Response 4:

Interactions between hydrological and N state variables are explained in section 2.4.1, pages 5676-5678.

Our initial paper was much longer, and we shortened it to about 10,000 words, which is still much longer than most manuscripts in *Biogeosciences*. Because this paper focuses on model development, we could not further shorten the model description sections, which led to a relatively short result section. At this point we prefer not to increase the length of the manuscript.

Reviewer's Comment 5:

P5671 L25. "Global", in which sense: planet scale or simulating all processes or both? I think authors are thinking for the spatial scale, but the multi-process aspect can be also important due to potential interactions between different state variables. See my comment concerning section 7.

Response 5:

To clarify the scale issue, will modify page 5671, lines 23-27 to read:

To characterize implications of human and climate driven perturbation in the earth N cycling and its implication for water and air quality, the next-generation of N cycling models need to (1) account for regional and local changes in terrestrial and aquatic ecosystem structure and functioning, (2) represent in a consistent manner emissions and transformation of N to air, rivers and coasts, and (3) be global in extent and integrated with climate and earth system models.

Reviewer's Comment 6:

P5675 L1-3. Can you explain better? In particular, how to link "historical reconstruction" with "land use change scenarios"? The same for "unique disturbance histories" in L7.

Response 6:

We propose to modify page 5675, lines 1-3 to read:

The model tracks hundreds of years of land use change using global land use transition scenarios that were historically reconstructed by combining satellite-based contemporary patterns of agriculture with historical data on agriculture and population (Hurt et al., 2006).

Furthermore, we will modify page 5675, lines 5-7:

The model is spatially distributed, and each grid cell consists of up to 15 tiles: 1 natural vegetation, 1 cropland, 1 pasture, and 1 to 12 secondary vegetation tiles representing unique disturbance histories (i.e. de/reforestation, agricultural practice change).

Reviewer's Comment 7:

P5672 and P5674: "vetetation"

Response 7:

Agreed, "vetetation" will be changed to "vegetation".

Reviewer's Comment 8:

P5677 and others. Add a sentence to introduce the equations.

Response 8:

We will add sentences introducing each of the equations:

page 5677, line 19:

Dissolved organic, ammonium, and nitrate N leaching from the soil are described as:

page 5678, lines 19-22:

Because soil nitrate contents are relatively low and limiting under natural conditions, we used a first-order loss function with respect to soil nitrate N content, with adjustments for the influence of soil water content and temperature to simulate soil denitrification rate:

page 5681, lines 7-8:

The N loads in a reach are routed downstream with the water as following.

Reviewer's Comment 9:

If authors like structured conclusions, they can be grouped into model characteristics, implementation results and exploitation.

Response 9:

Agreed, we will change the order of the conclusions by moving a sentence in page 5690, lines 11-14 to between line 19 and 20.

<A.J. Dolman (Referee)>

Reviewer's Comment 1:

p5677 Storage in groundwater, this is issue as the authors note, given the 12-80% additional delivery rate. Can the authors comment on the need for further model development in this area in the discussion? It seems that calibrating this for every catchment in the world is a bit of an issue.

Response 1:

Agreed, we propose to modify page 5677, lines 18-19 to read:

The need to incorporate these calibration factors, which are at the present basin specific, indicates that future improvements to LM3-TAN should focus on resolving these processes (i.e. N cycle in microbes, reservoirs, and vertically distributed soil layers).

Reviewer's Comments 2 & 3:

p5686 Validation. the use of a correlation coefficient is a bit poor. Is it possible to have some more advanced metrics, like the Nash Sutcliffe efficiency that are more common in hydrological models?

p5686 Is it possible to add a table with these results. This makes a comparison between the various outputs a lot easier.

Response 2 & 3:

In addition to R^2 , we will present two more correlation coefficients (Pearson and Spearman) with the corresponding p-values in a new table (Table 5).

Reviewer's Comments 4:

p5686 I am somewhat surprised by the low r^2 of the discharge values and how the 28% lower flows translate back into the N-transport uncertainty.

Response 4:

The reason for the low r^2 is a bias in the flow predictions during wet years. To address the low r^2 we will modify page 5686, lines 7-12 explaining:

Using global hydrological data and a universal parameter set for the entire watershed, the model produced reasonable temporal patterns of annual stream discharge. The simulated stream discharges were in a good agreement with the reported values in dry years and periods (July to September), but under-estimated stream discharges in wet years and periods (March to May).

Overall, although the 19-year average simulated discharge was about 28% lower than the corresponding reported value, their linear and rank correlations were significantly high (Table 5), implying that the bias was systemic and accounted for in the calibration of the N species.

Reviewer's Comments 5:

p5701 Figure 1 is rather complicated to read with too many arrow crossing the boxes. Either simplify or redraw to make it more clear (i.e. "route" the arrow along the boxes).

Response 5:

We fully agree, however, we tried many ways to draw this figure, and as presented it is the simplest way not to lose important model components. Routing the arrows along the boxes increase lines and makes the figure more difficult to interpret. Our preference is to leave Fig. 1 as is. We made the dashed arrow lines clearer.

<Q. Zhang>

Reviewer's Comments 1:

Table 3: For the 9 short-term SRBC sites, it seems that only limited periods of data (around 1 year or less) were used in the model evaluation, given that the modeling period ends in 2005. Therefore, the author may consider specifying how many (or which) months of data at these new sites were used.

Response 1:

In this research, we only used "annual" N loads to evaluate the model. Fig. 5 shows annual N loads at the Marietta station for the period 1987-2005. Fig. 6 shows 17 year (1989-2005) average N loads for the 6 long-term stations and annual N loads for the year 2005 for the 15 monitoring stations including the 9 short-term stations.

Reviewer's Comments 2:

Table 4: It is not very clear how the non-point and point source loads for each subbasins were obtained. If this is from literature or SRBC, a citation would help readers better understand the source.

Response 2:

We cited all of the used non-point and point N sources for the entire basin in the section 5, page 5684, lines 1-21.

Reviewer's Comments 3:

Figure 3: This is a comprehensive map. Nicely done. However, what is missing from this map is the Conowingo site (managed by USGS) at the basin outlet. It is surprising (and also a waste of resource) that the Conowingo site was not used in the model, which has even longer N data for multiple sub-species than any of the long-term SRBC sites (1979). To clarify, Marietta is actually the most downstream "SRBC" site. It may be misleading if not including "SRBC" (in line 5 of Section 7.1). To me, it would be very interesting to see how much the results would be affected by including this outlet site in the model.

If the authors are concerned with the reservoir system located between Marietta and Conowingo sites and want to exclude potential influences by the reservoir system on their model, then such kind of considerations should be at least briefly discussed in the text. In fact, N load has not been affected much by the reservoir system, at least not to the extent that phosphorous and sediment have. These issues have been recently examined by Hirsch (2012)

[<http://pubs.usgs.gov/sir/2012/5185/>] and Zhang et al. (2013)

[<http://www.sciencedirect.com/science/article/pii/S0048969713001757>].

Response 3:

Agreed, we will modify "the last downstream station Marietta" to "the last downstream SRBC station Marietta" in page 5686, line 5 and page 5670, lines 17-18.

We will modify Table 3 to include the Conowingo site description.

We will modify Figure 3 to include the Conowingo site on the map.

We will modify Figure 5 to include simulated and reported nitrate-N loads at Conowingo.

We will modify “15 monitoring stations” to “16 monitoring stations” in page 5670, line 14; page 5674, line 2; and page 5686, line 4.

We will modify page 5683, lines 7-13 to read:

Chemical constituents of the basin’s water were monitored by the USGS and Susquehanna River Basin Commission (SRBC). One USGS and six SRBC long-term nutrient monitoring sites monitored since 1985 and 9 newly introduced SRBC sites monitored since either 2004 or 2005 to present (Table 3; Fig. 3; McGonigal, 2011; USGS, 2014) were chosen for model evaluation. The 16 sites vary in sub-basin area and land use. Among the USGS and SRBC sites, the Conowingo and Marietta sites on the main channel of the Susquehanna River have the largest sub-basin areas respectively (70,189 and 67,314 km²).

We will modify “(ESTIMATOR; SRBC)” to “(ESTIMATOR; SRBC, 2006; USGS, 2014)” in page 5683, line 22.

This is a valid comment and we plan to modify Page 5687, lines 3-5 to read:

The model also produced reasonable temporal patterns of annual dissolved-N ($r = 0.7$), nitrate-N ($r = 0.6$), ammonium-N ($r = 0.7$), and dissolved organic-N ($r = 0.6$) loads (Fig. 5; Table 5). At Conowingo, 20 year average simulated nitrate-N load agreed well with the corresponding reported value (-3.7%), but the model, which doesn't have lakes or reservoirs, fails to capture inter-annual variations of the loads ($r = 0.2$), which are affected by the reservoir system between the Marietta and Conowingo monitoring sites (Fig. 5).

Because Marietta is very close to Conowingo and its sub-basin covers 95 percent of the entire Susquehanna watershed, calibrating the model based on the reported data at Marietta does not make a significant difference. Still, we agree that Conowingo is a better site for model calibration since it is the last downstream station. We will use the data from the Conowingo site when we calibrate our next version of the model capable of simulating N in lakes and reservoirs.

Reviewer’s Comments 4:

Figure 4: The figure caption is a little confusing for (c) and (d) – do they only differ by unit? Why "fertilizer, manure, and legume applications" are grouped? In addition, unit is missing for CSO (b).

Response 4:

We will add the unit (kg km⁻² yr⁻¹) for the CSO in Fig. 4.

We designed the model to apply agricultural N sources (fertilizer, manure, and legume applications) into crop land use. Section 2.1 (page 5675, lines 1-7), section 5 (page 5685 lines 3-7), and Fig. 1 may help to understand the model structure. We have four different land uses in a grid cell. Fig. 4c shows the agricultural N application rate to each grid cell, while Fig. 4d shows the agricultural N application rate to each crop land use in each grid cell.

Reviewer's Comments 5:

Figure 5: The word "observed data" is not strictly correct, since the SRBC loads were calculated based on the ESTIMATOR model. I suggest alternative wording such as "estimated load based on monitoring data", if a definition of "observed" is not given in the caption.

Response 5:

"observed" will be changed to "reported".

Reviewer's Comments 6:

Figure 6: It might be more clear if two contrasting colors are used, instead of light and dark colors.

Response 6:

We prefer the figure as is since it is clear if it is printed in black and white. We assume that many readers will print the document in B&E.

Reviewer's Comments 7:

Figure 7: It is inconsistent with other figures that here the unit is in logarithm scale. If this is because the values are rather large, why not use units such as " 10^6 kg/year"?

Response 7:

For Fig. 7, we used a logarithmic scale because the results cover a large range of values. This way helps the small stream N loads to be mapped in different colors. In fact, using actual values instead of the logarithms results in figures very similar to Fig. 10 where most of values range from 0-2.5% and are described in only one color, purple. Hence, we prefer to keep the logarithmic scale for this figure.

<Author Comment>

Comment 1

We will clarify the capacity of LM3-TAN to describe N dynamics with a single universal parameter set for the study of Susquehanna watershed and potential limitation of LM3-TAN when applying it to other watersheds very different from the Susquehanna watershed. We also will suggest how LM3-TAN may be able to meet the challenge in the discussion.

We will modify page 5672, lines 16-21 to read:

In addition, many regional models (e.g. EPIC, ANIMO), which have been applied to far smaller basins compared to the Susquehanna watershed, often use basin-specific parameters for mineralization, nitrification, and denitrification, which complicates their application on a global scale for decadal-to-century scale studies. LM3-TAN is capable of describing N dynamics with a universal parameter set – the same parameters for all of the sub-basins within an area of 71,220 km² and time periods for this simulation.

We will add the following sentences to page 5680, line 15:

Still, care has to be taken when applying the model to other watersheds that may be very different in terms of soil and climate properties from the Susquehanna watershed. Furthermore, since soil denitrification becomes zero-order in extreme nitrate rich environment, instead of using the first-order loss function for all of the land use types, using a Monod function for agricultural land use may help LM3-TAN's global application where N loadings would vary widely.

Comment 2

We will explain an implication of the size of a calibration factor by adding a sentence in page 5677, line 18.

Considering its importance in groundwater, a relatively larger size of the nitrate N factor is expected.

Comment 3

We will explain how river depth has influence on denitrification rate by adding sentences in page 5680, line 25.

River denitrification happens mainly in the benthic and/or hyporheic zones. Therefore, a river denitrification rate that is inversely proportional to the river depth accounts for the ratio of water column to benthic area.

LIST OF ALL RELEVANT CHANGES MADE IN THE MANUSCRIPT

Page 5669

A co-author was added:

P. C. D. Milly³

[3]{U. S. Geological Survey and NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA}

Page 5670, lines 17-18

“the last downstream station Marietta” to “the last downstream SRBC station Marietta”

Page 5670, lines 14 & Page 5674, lines 2

“15 monitoring stations” to “16 monitoring stations”

Page 5670, lines 22

“contribute” to “contributed”

Page 5670, lines 24

“the soil denitrification rates” to “soil denitrification rates”

Page 5670, lines 25 and 26, Page 5687, lines 27, & Page 5690, lines 23

“sub-basin” to “Sub-basin”

Page 5671, lines 23-27

“To characterize ~ and coasts.” to

“To characterize implications of human and climate driven perturbation in the earth N cycling and its implication for water and air quality, the next-generation of N cycling models need to (1) account for regional and local changes in terrestrial and aquatic ecosystem structure and functioning, (2) represent in a consistent manner emissions and transformation of N to air, rivers and coasts, and (3) be global in extent and integrated with climate and earth system models.”

Page 5672, lines 3-5

“varies in ~ of observations” to

“whose sub-basins vary in climate, land use, and associated N sources and transformations, with the detailed dataset of observations.”

Page 5674, line 8, Page 5677, line 15, Page 5679, line 23, Page 5685, line 2 & 21, & Page 5687, line 11

“vetetation” to “vegetation”

Page 5672, lines 16-21

“In addition, ~ time periods.” to

“In addition, many regional models (e.g. EPIC, ANIMO), which have been applied to far smaller basins compared to the Susquehanna watershed, often use basin-specific parameters for mineralization, nitrification, and denitrification, which complicates their application on a global scale for decadal-to-century scale studies. LM3-TAN is capable of describing N dynamics with a universal parameter set – the same parameters for all of the sub-basins within an area of 71,220 km² and time periods for this simulation.”

Page 5672, line 28 & Page 5674, line 16

“observed” to “reported”

Page 5675, lines 1-3

“The model ~ (Hurt et al., 2006).” to

“The model tracks hundreds of years of land use change using global land use transition scenarios that were historically reconstructed by combining satellite-based contemporary patterns of agriculture with historical data on agriculture and population (Hurt et al., 2006).”

Page 5675, lines 5-7

“The model ~ disturbance histories.” to

“The model is spatially distributed, and each grid cell consists of up to 15 tiles: 1 natural vegetation, 1 cropland, 1 pasture, and 1 to 12 secondary vegetation tiles representing unique disturbance histories (i.e. de/reforestation, agricultural practice change).”

Page 5675, lines 23-24

“fire and hydrological leaching” to ” fire, hydrological leaching, and mineralization”

Page 5677, lines 18-19

“The need ~ these processes” to

“Considering its importance in groundwater, a relatively larger size of the nitrate N factor is expected. The need to incorporate these calibration factors, which are at the present basin specific, indicates that future improvements to LM3-TAN should focus on resolving these processes (i.e. N cycle in microbes, reservoirs, and vertically distributed soil layers). Dissolved organic, ammonium, and nitrate N leaching from the soil are described as:”

Page 5678, lines 19-22

“Because soil ~ and temperature.” to

“Because soil nitrate contents are relatively low and limiting under natural conditions, we used a first-order loss function with respect to soil nitrate N content, with adjustments for the influence of soil water content and temperature to simulate soil denitrification rate:”

Page 5680, line 15

Added sentences after “sub-basins.”

“Still, care has to be taken when applying the model to other watersheds that may be very different in terms of soil and climate properties from the Susquehanna watershed. Furthermore, because soil denitrification becomes zero-order in extreme nitrate rich environment, instead of using the first-order loss function for all of the land use types, using a Monod function for agricultural land use may help LM3-TAN’s global application where N loadings would vary widely.”

Page 5680, line 25

Added sentences after “reach (Alexander et al., 2009)”

“River denitrification happens mainly in the benthic and/or hyporheic zones. Therefore, a river denitrification rate that is inversely proportional to the river depth accounts for the ratio of water column to benthic area.”

Page 5681, lines 7-8

“The N ~ the water” to

“The N loads in a reach are routed downstream with the water as following.”

Page 5682, lines 22

“classic” to ”clastic”

Page 5683, lines 5

“gages” to ”gauges”

Page 5683, lines 7-13

“Chemical constituents ~ (67,314 km²)” to

“Chemical constituents of the basin’s water were monitored by the USGS and Susquehanna River Basin Commission (SRBC). One USGS and six SRBC long-term nutrient monitoring sites monitored since 1985 and 9 newly introduced SRBC sites monitored since either 2004 or 2005 to present (Table 3; Fig. 3; McGonigal, 2011; USGS, 2014) were chosen for model evaluation. The 16 sites vary in sub-basin area and land use. Among the USGS and SRBC sites, the Conowingo and Marietta sites on the main channel of the Susquehanna River have the largest sub-basin areas respectively (70,189 and 67,314 km²).”

Page 5683, lines 22

“(ESTIMATOR; SRBC)” to “(ESTIMATOR; SRBC, 2006; USGS, 2014)”

Page 5685, lines 24-25

“We then ~ to 2005” to

“We then applied the reported 1985’s anthropogenic N data from 1954 to 1984, and reported annual anthropogenic N data from 1985 to 2005.”

Page 5686, lines 3-7

“Stream discharges ~ be assessed” to

“We simulated with LM3-TAN stream dissolved organic-N, ammonium-N, and nitrate-N loads throughout the river network. The model was calibrated by comparing the modeled stream N loads with the corresponding reported N loads at the last downstream SRBC station Marietta, in which contributions of the entire watershed to the stream flows and N loads can be assessed. Thus, temporal evaluation of the stream discharges and N loads for the period 1987-2005 was focused on at the Marietta station. River data from the 16 monitoring stations (1986-2005) were also used to evaluate spatial stream discharges and N loads.”

Page 5686, lines 7-12

“In general ~ reported values.” to

“Using global hydrological data and a universal parameter set for the entire watershed, the model produced reasonable temporal patterns of annual stream discharge. The simulated stream discharges were in a good agreement with the reported values in dry years and periods (July to September), but under-estimated stream discharges in wet years and periods (March to May). Overall, although the 19-year average simulated discharge was about 28% lower than the corresponding reported value, their linear and rank correlations were significantly high (Table 5), implying that the bias was systemic and accounted for in the calibration of the N species.”

Page 5686, lines 17

“nitrate N ($R^2 = 0.98$)” to “nitrate N load ($R^2 = 0.98$)”

Page 5686, lines 18

“for ammonium N ($R^2 = 0.85$)” to “for ammonium N ($R^2 = 0.85$) loads”

Page 5687, lines 3-5

“The model ~ loads (Fig. 5).” to

“The model also produced reasonable temporal patterns of annual dissolved-N ($r = 0.7$), nitrate-N ($r = 0.6$), ammonium-N ($r = 0.7$), and dissolved organic-N ($r = 0.6$) loads (Fig. 5; Table 5). At Conowingo, 20 year average simulated nitrate-N load agreed well with the corresponding reported value (-3.7%), but the model, which doesn't have lakes or reservoirs, fails to capture inter-annual variations of the loads ($r = 0.2$), which are affected by the reservoir system between the Marietta and Conowingo monitoring sites (Fig. 5).”

Page 5687, lines 16

“Efficiency” to “ability”

Page 5689, line 21

“in between” to “between”

Page 5690, lines 11-19

The order of the conclusions was changed by moving a sentence in page 5690, lines 11-14 to between line 19 and 20.

Page 5690, lines 13

“resolving” to “resolves”

Page 5690, lines 20

“is sensitive the most” to “is most sensitive”

Page 5690, lines 27

“Mariette“ to “Marietta“

Page 5691, lines 8-9

“thank P. C. D. Milly and Krista A. Dunne of the“ to “ thank Krista A. Dunne from the“
“hydrologic module“ to “the hydrologic module“

Page 5696, between lines 10 and 11

Added sentences

“We thank Joel Blomquist from the United States Geological Survey for assisting with interpretation of N load data at the USGS Conowingo sampling station. We thank Kevin McGonigal from the Susquehanna River Basin Commission for providing us with the various N load data.”

Page 5691, lines 19

“Bohlke“ to “Böhlke“

Page 5694, lines 15

“in review” to “doi: <http://dx.doi.org/10.1175/JHM-D-13-0162.1>”

Page 5696, between lines 14 and 15

Added a reference “USGS (US Geological Survey): available at:
http://cbrim.er.usgs.gov/loads_query.html (last access: 28 July 2014), 2014.”

Table 1

- Merged columns in the row “Reference or Rationale” misled like that some of parameters did not have any reference or rationale, which is not true. So, I did not merge any columns in the “Reference or Rationale” row for different parameters which have the same reference or rationale.
- Other modifications were made in the following columns and rows (column, row).
(4, 5), (6, 5), & (21, 5): “fit” to “calibrated”
(7, 2) & (8, 2): “Transfer” to “transfer”

Table 2

- Table 2 caption was modified to:
“Definition of prognostic (PV) and diagnostic (DV) variables and inputs/forcings (IF) used in the equations.”
- A row was added to specify if the variables are prognostic or diagnostic or inputs/forcings.
- First two columns were merged and the contents were modified to:
“Vegetation and Soil Equations”
- The content in the 16th column was modified to:
“River Equations”

Table 3

- Table 3 caption was modified to:
“Susquehanna River Basin Geographic Statistics for the USGS and SRBC nutrient monitoring sites (McGonigal, 2011; USGS, 2014).”
- A column was added to include the Conowingo site description.
- “6 Long-term Sites” to “7 Long-term Sites”

Table 5

A new table 5 was added.

Figure 1

Figure 1 was changed.

Figure 3

- Figure 3 was changed.
- Figure 3 caption was modified to:
“Map of the Susquehanna watershed, showing 6 major sub-basins, main stem of the Susquehanna River, major tributaries (Chemung, West Branch Susquehanna, and Juniata River), streams, and the location of USGS stream gauges and USGS and SRBC nutrient monitoring sites.”

Figure 4

- Figure 4 caption was modified.
“combined sewer overflow (b),” to “combined sewer overflow (kg/km² year) (b),”

Figure 5

- Figure 5 was changed.
- Figure 5 caption was modified to:
“20 years (1986-2005) of the simulated stream N loads (normalized by sub-basin areas) at Marietta and Conowingo and the corresponding reported data from SRBC and USGS.”

Figure 9

Figure 9 was changed.