Assessing landscape N removal in coastal New England catchments using the N-Sink approach with a new R package, nsink

Hollister. J. W. \* *1*, Kellogg, D. Q. *2*, Arnold, C.*3*, Gold, A.*2*, Wilson, E. H.*3*, Chadwick, C.*3*, Dickson, D.\**3*, Forshay, K. J.*4*

*1* US Environmental Protection Agency, Office Of Research and Development, Atlantic Coastal Environmental Sciences Division, Narragansett, RI 02882

*2* University of Rhode Island, Department of Natural Resources Science, Kingston, RI 02881

*3* University of Connecticut, Center for Land Use Education and Research, Haddam, CT 06438

*4* US Environmental Protection Agency, Office Of Research and Development, Groundwater Characterization and Remediation Division, Ada, OK 74820

\* \*corresponding author: [qkellogg@uri.edu\*](mailto:qkellogg@uri.edu*)

Excessive nitrogen (N) loading to coastal ecosystems impairs estuarine water quality. Land management decisions made within estuarine watersheds have a direct impact on downstream N delivery. Natural features within watersheds can act as landscape sinks for N – wetlands, streams and ponds – transforming dissolved N into gaseous N, effectively removing it from the aquatic system. Identifying and evaluating these landscape sinks and their spatial relationship to N sources can assist managers to evaluate the effects of their decisions on downstream resources. N-Sink is an approach that uses widely available GIS data to identify landscape sinks within HUC-12 catchments, estimate their N removal potential and map the effect of those sinks on N movement through the catchment. A series of static maps are produced to visualize N removal efficiency, N loading, transport and delivery, the latter three in the form of indices, To facilitate data acquisition, processing and visualization, an R package nsink was developed and has been used to produce static maps for all HUC-12 catchments in southern New England. Users can use the R package to investigate a HUC-12 of interest, or can visit the UConn website to explore the already mapped areas. Users can also investigate specific flowpaths interactively by clicking on any location within the catchment. A flowpath is generated, along with a table describing N removal along each segment.

# Introduction

Excess nitrogen (N) delivery via surface water to coastal estuaries contributes to impaired water quality evidenced by excess algal blooms and hypoxia (Ryther and Dunstan 1971, Nixon 1995, Howarth and Marino 2006). Hydrologic connections and flowpaths are influential in the delivery of nutrients to estuaries (Mengistu et al. 2020). Identifying landscape N sinks along hydrologic flowpaths – areas where dissolved N is transformed to gaseous N and effectively removed from the aquatic system – and their effect on N delivery at the watershed scale is helpful to watershed managers, land use planners and conservation organizations. The intent of N-Sink, and its associated R package nsink, is as a screening tool to identify areas where downgradient N sinks are able to process excess N vs. areas where downgradient N sinks are less prevalent or effective at removing N. The latter argues for more aggressive N management at the source, while the former argues for protecting and/or restoring downgradient N sinks that are helping to reduce N delivery to estuaries. N-Sink was developed as a web-based tool to visualize and explore landscape N sinks at the HUC-12 scale, and makes extensive use of widely available GIS data. The theoretical underpinnings are outlined in (Kellogg et al. 2010) and build on peer-reviewed meta-analyses and reviews (Seitzinger et al. 2006, Alexander et al. 2007, i.e., Mayer et al. 2007) to estimate N removal within landscape N sinks – wetlands, streams and ponds. Residence time in these landscape features is the primary driver of N retention and transformation (e.g., Klocker et al. 2009).

The nsink R package was written to simplify the acquisition and management of the spatial data necessary to estimate N removal within each identified landscape sink, estimate cumulative N removal along a specified flowpath, and create a set of static watershed maps. These static maps include 1) N Removal Efficiency, a color-coded map of the estimated N removal efficiency of the different landscape sinks; 2) N Transport Index, a heat map with the cumulative N removal along flowpaths originating from a grid of points across a watershed, highlighting “leaky” areas with less downstream N retention vs. those with higher downstream retention; and c) N Delivery Index, combining maps (a) and (b) along with source loading based on NLCD categories and expressed as an N index ranging from 0 to 1, resulting in a map showing potential N delivery from different sources, taking into account the potential N removal as water moves downstream.

# Methods

## Development of the R package nsink

The package nsink implements an approach to estimate relative nitrogen (N) removal along a flow path. This approach is detailed in Kellogg et al. (2010). The nsink package follows from an initial version of N-Sink written in ArcGIS using ModelBuilder. That version required time-consuming data manipulation by hand due to limitations of earlier NHD datasets as well as the limitations of working in a GIS environment requiring a user license. With the increasingly more versatile GIS packages now available in R, the previously time-consuming spatial data acquisition and management of N-Sink could be automated and more easily applied to other HUC-12 catchments, leading to the development of the nsink package. Specifically, nsink relies on packages sf and raster.

## Using the nsink package

The nsink package provides several functions to set up and run an N-Sink analysis for a specified 12-digit HUC code. All required data are downloaded, prepared for the analysis, HUC-wide nitrogen removal calculated, and flow paths summarized. Additionally, a convenience function that will run all of the required functions for a specified HUC is included.

The workflow follows a simple series of steps to create a set of static maps: N Removal Efficiency, N Transport Efficiency, and N Delivery Index.

## Download the data

nsink\_get\_huc\_id() - search all 12-digit HUC names using a known location name.

nsink\_get\_data() - download and save in a data folder hydrography, soils, and land cover for the specified 12-digit HUC. These data will cover the HUC, but will not yet be reduced to those data exclusive to the specified HUC. The nsink package utilizes openly available data from several U.S. Federal Government sources. All data are from publicly available sources and as of 2021-06-17 no authentication is required to access these sources. The HUC ID is required and users may specify a path for storing the data as well as indicate whether or not to download the data again if they already exist in the data directory.

## Prepare the data

Once the data are downloaded there are several additional data processing steps that are required to subset the data to the HUC and set all data to a common coordinate reference system (CRS).

These include:

* filter out the HUC boundary
* mask all other data sets to the HUC boundary
* convert all column names to lower case
* create new columns
* harmonize raster extents
* set all data to common CRS

The nsink\_prep\_data() function will complete all of these processing steps.  
It requires a HUC ID, a specified CRS, and a path to a data directory. It returns a “list” (a specific data format in R) with all required data for subsequent N-Sink analyses. NOTE: Not sure where we are at with the CRS. Check for latest. A quick note on the CRS. In the near future, the preferred way to specify the CRS values will either be with Well-Known Text (WKT) or [EPSG Codes](http://spatialreference.org/ref/epsg/). Proj.4 strings will eventually be deprecated. Currently the packages that nsink relies on are a t different stages in implementing the changes to PROJ. Currently Proj.4 strings are preferred, but that will change soon.

## Calculate N removal

The next step in the N-Sink process is to calculate relative nitrogen removal.  
Details on how the nitrogen removal estimates are calculated are available in Kellogg et al. (2010). Since the publication of that article, changes have been implemented to make better use of the most recently available geospatial data; data that were not available at the time of publication. Those changes are described in Section XX. The nsink\_calc\_removal() function takes the prepared data as an input and returns a “list” with three items:

* raster\_method: This item contains a raster based approach to calculating removal. Used for the static maps of removal.
* land\_removal: This represents land based nitrogen removal which is hydric soils with areas of impervious surface removed.
* network\_removal: This contains removal along the NHD Plus flow network.  
  Removal is calculated separately for streams and waterbodies (i.e., lakes and reservoirs).

## Generate and summarize flowpaths

A useful part of the N-Sink approach is not just the calculation of relative removal for individual components of the landscape, it is the ability to summarize that removal along the legnth of a specified flowpath. The nsink package provides two functions that facilitate this process. The nsink\_generate\_flowpath() function takes a point location as an sf object and the prepped data (generated by nsink\_prep\_data()) as input and returns an sf LINESTRING of the flowpath starting from the input point location and terminating at the furthest downstream location in the input NHD Plus. The flowpath on land is generated from a flow direction grid. Once that flowpath intersects the surface water network, flow is determined by flow along the NHD Plus stream network. First, create the sf POINT object.

# Load up the sf package  
library(sf)

## Linking to GEOS 3.8.1, GDAL 3.1.4, PROJ 6.3.1

# Starting point  
pt <- c(1948121, 2295822)  
start\_loc <- st\_sf(st\_sfc(st\_point(c(pt)), crs = aea))

You may also determine your point location interactively by plotting your data and using the locator() function . First create a simple plot.

# Create a simple plot  
plot(st\_geometry(niantic\_data$huc))  
plot(st\_geometry(niantic\_data$lakes), add = T, col = "darkblue")  
plot(st\_geometry(niantic\_data$streams), add = T, col = "blue")

With the map made, you can use that to interactively select a location and use the x and y to create the sf POINT object.

# Select location on map for starting point  
pt <- unlist(locator(n = 1))  
# convert to sf POINT  
start\_loc\_inter <- st\_sf(st\_sfc(st\_point(pt), crs = aea))

With a point identified, we can use that as the starting location for our flowpath.

niantic\_fp <- nsink\_generate\_flowpath(start\_loc, niantic\_data)

The reutrned value has both the flowpath\_ends, the portion of the flowpath on the land which is created using the flow direction grid, and the flowpath\_network which is the portion of the flowpath from the NHD Plus network that occur after the upstream flowpath\_ends intersect the surface water network.

niantic\_fp

## $flowpath\_ends  
## Simple feature collection with 2 features and 0 fields  
## Geometry type: LINESTRING  
## Dimension: XY  
## Bounding box: xmin: 1948003 ymin: 2281809 xmax: 1958233 ymax: 2296449  
## CRS: +proj=aea +lat\_0=23 +lon\_0=-96 +lat\_1=29.5 +lat\_2=45.5 +x\_0=0 +y\_0=0 +ellps=GRS80 +towgs84=0,0,0,0,0,0,0 +units=m +no\_defs  
## fp\_ends  
## 1 LINESTRING (1948123 2295819...  
## 2 LINESTRING (1957513 2283579...  
##   
## $flowpath\_network  
## Simple feature collection with 30 features and 18 fields  
## Geometry type: LINESTRING  
## Dimension: XY  
## Bounding box: xmin: 1949097 ymin: 2283213 xmax: 1957637 ymax: 2296345  
## CRS: +proj=aea +lat\_0=23 +lon\_0=-96 +lat\_1=29.5 +lat\_2=45.5 +x\_0=0 +y\_0=0 +ellps=GRS80 +towgs84=0,0,0,0,0,0,0 +units=m +no\_defs  
## First 10 features:  
## stream\_comid fdate resolution gnis\_id gnis\_name lengthkm  
## 1 6170640 2008-07-17 Medium 208394 Latimer Brook 0.658  
## 2 6170090 1999-09-06 Medium 208394 Latimer Brook 0.130  
## 3 6170644 2008-07-17 Medium 208394 Latimer Brook 0.307  
## 4 6170096 1999-09-06 Medium 208394 Latimer Brook 1.241  
## 5 6170966 2008-07-17 Medium 208394 Latimer Brook 0.148  
## 6 6170652 2008-07-17 Medium 208394 Latimer Brook 0.169  
## 7 6170100 1999-09-06 Medium 208394 Latimer Brook 0.071  
## 8 6170658 2008-07-17 Medium 208394 Latimer Brook 0.268  
## 9 6170664 2008-07-17 Medium 208394 Latimer Brook 0.290  
## 10 6170122 1999-09-06 Medium 208394 Latimer Brook 1.212  
## reachcode flowdir lake\_comid ftype fcode shape\_leng  
## 1 01100003000098 With Digitized 6169112 ArtificialPath 55800 0.006702425  
## 2 01100003000097 With Digitized 0 StreamRiver 46006 0.001281976  
## 3 01100003000097 With Digitized 6169134 ArtificialPath 55800 0.003591168  
## 4 01100003000097 With Digitized 0 StreamRiver 46006 0.013000742  
## 5 01100003000097 With Digitized 6169152 ArtificialPath 55800 0.001734513  
## 6 01100003000096 With Digitized 6169152 ArtificialPath 55800 0.001765153  
## 7 01100003000096 With Digitized 0 StreamRiver 46006 0.000853676  
## 8 01100003000796 With Digitized 6169150 ArtificialPath 55800 0.002886813  
## 9 01100003000797 With Digitized 6169150 ArtificialPath 55800 0.002836633  
## 10 01100003000094 With Digitized 0 StreamRiver 46006 0.012087397  
## enabled gnis\_nbr totma fromnode tonode stream\_order  
## 1 True 0 257.19887393 150029595 150029459 1  
## 2 True 0 0.01071115 150029459 150029597 1  
## 3 True 0 12.68401556 150029597 150029461 1  
## 4 True 0 0.05086744 150029461 150029733 1  
## 5 True 0 1.29543583 150029733 150029601 1  
## 6 True 0 1.47924767 150029601 150029463 2  
## 7 True 0 0.00302952 150029463 150029604 2  
## 8 True 0 8.76143527 150029604 150029607 2  
## 9 True 0 9.48065757 150029607 150029468 2  
## 10 True 0 0.05996156 150029468 150029471 2  
## geometry  
## 1 LINESTRING (1949097 2296345...  
## 2 LINESTRING (1949129 2296257...  
## 3 LINESTRING (1949211 2296168...  
## 4 LINESTRING (1949488 2296233...  
## 5 LINESTRING (1950416 2295753...  
## 6 LINESTRING (1950556 2295764...  
## 7 LINESTRING (1950689 2295680...  
## 8 LINESTRING (1950756 2295705...  
## 9 LINESTRING (1950974 2295603...  
## 10 LINESTRING (1950975 2295347...

With a flowpath generated we can summarize the relative nitrogen removal along that flowpath with the nsink\_summarize\_flowpath() function. It takes the flowpath and removal as input. A data frame is returned with each segment identified by type, the percent removal associated with that segment, and relative removal. Total relative removal is 100 - minimum of the n\_out column. That is, n\_out keeps track of the proportion of N leaving each subsequent downstream segment.

niantic\_fp\_removal <- nsink\_summarize\_flowpath(niantic\_fp, niantic\_removal)  
niantic\_fp\_removal

## # A tibble: 26 x 5  
## segment\_type length\_meters percent\_removal n\_in n\_out  
## <chr> <dbl> <dbl> <dbl> <dbl>  
## 1 Hydric 154 2.56 100 97.4  
## 2 No Removal 1418 0 97.4 97.4  
## 3 Lake/Pond 100 54.8 97.4 44   
## 4 Stream 130 0.014 44 44   
## 5 Lake/Pond 307 18.8 44 35.7  
## 6 Stream 1241 0.051 35.7 35.7  
## 7 Lake/Pond 317 6.71 35.7 33.3  
## 8 Stream 71 0.003 33.3 33.3  
## 9 Lake/Pond 558 34.1 33.3 22   
## 10 Stream 1212 0.039 22 22   
## # … with 16 more rows

100-min(niantic\_fp\_removal$n\_out)

## [1] 78.1

## Static maps

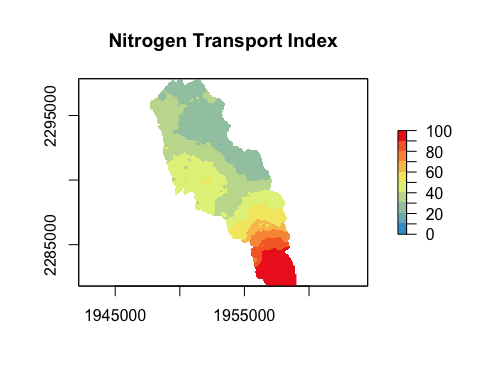
Individual flow paths are useful for specific areas of interest, but it is also useful to look at removal patterns across the landscape. The nsink\_generate\_static\_maps() function provides these HUC-wide rasters. Required inputs are the prepped data, removal raster, and sampling density, for which a default is provided.  
The function returns four separate rasters.

* removal\_effic: Landscape wide estimate of relative nitrogen removal as a percentage.
* loading\_idx: An index of relative nitrogen loads by land cover class derived from published sources.
* transport\_idx: This is relative nitrogen transport for a sample of all possible flowpaths in a given HUC. This is an expensive computational task, so nsink generates a removal hotspot map based on a sample of flowpaths and the final hotspot map is interpolated from these samples and referred to as the nitrogen transport index. The samp\_density argument controls the number of sample flowpaths generated.
* delivery\_idx: The delivery index is the combination of the loading index and the transport index It represents which areas of the landscape are delivering the most nitrogen to the outflow of the watershed.

niantic\_static\_maps <- nsink\_generate\_static\_maps(niantic\_data, niantic\_removal,   
 900)

And with these static maps generated, you can plot them quickly with nsink\_plot\_removal(), nsink\_plot\_delivery(), or nsink\_plot\_transport().  
An example of nsink\_plot\_transport() is below.

nsink\_plot\_transport(niantic\_static\_maps$transport\_idx)



## Convenience function: Build it all!

The workflow described above includes all the basic functionality. Some users may wish to use nsink to calculate the base layers for an N-Sink analysis and then build an application outside of R. A convenience function that downloads all data, prepares, calculates removal, and generates static maps has been included to facilitate this type of analysis. The nsink\_build() function requires a HUC ID, coordinate reference system, and sampling density. An output folder is also needed but has a default location. There are ptional arguments for forcing a new download and playing a sound to signal when the build has finished. All prepped data files and .tif files are written to the output folder for use in other applications.

niantic\_huc <- nsink\_get\_huc\_id("Niantic River")$huc\_12  
aea <- "+proj=aea +lat\_1=29.5 +lat\_2=45.5 +lat\_0=23 +lon\_0=-96 +x\_0=0 +y\_0=0   
+ellps=GRS80 +towgs84=0,0,0,0,0,0,0 +units=m +no\_defs"  
nsink\_build(niantic\_huc, aea, samp\_dens = 900)

Currently we have created the static maps for all HUC-12s in coastal southern New England (map). These may be viewed and explored interactively at <https://clear.uconn.edu/projects/nsink/>.

## Changes to underlying methodology:

Several improvements have been made to the underlying methodology that have allowed nsink to make use of more recently available spatial data that had to be estimated in the original version. In estimating N removal in streams, we originally had to estimated stream depth and velocity based on other available geospatial data to arrive at a travel time along a given stream reach. With the updated and expanded NHD Plus V2 we can use the estimates provided in that dataset, making use of USGS expertise that went into these estimates.

In estimating N removal occurring in the terrestrial portion of the flowpath, we originally focused on riparian wetlands, using SSURGO mapped hydric soils to identify this landscape sink. With the latest version of SSURGO we include all hydric soils in the catchment, except those identified as “subaqueous.” Each soil mapping unit (SMU) that has a positive hydric status also has an estimate of %hydric for that SMU. We then estimate N removal within that SMU as 80% of of the %hydric. Each raster has one SMU identified with it and raster cell size is 30m X 30m, which is the maximum wetland width originally estimated as having 80% N removal. In the current version N removal is cumulative as a flowpath intercepts hydric soils in any cell before encountering surface water.

## Off-network hydrology

In the NHD Plus dataset there are situations where a water body or stream segment is “off-network,” i.e., it is not linked to the larger surface flow network. These may be features such as groundwater fed kettle ponds, or canals or ditches where connecting hydrology may not be captured in the NHD Plus dataset. These off-network features tend to have less NHD spatial data associated with them, making direct N removal estimates impossible. In these cases, we must estimate removal based on other in-network surface water features. While off-network features are rare and spatially small compared to the overall system, they do occur in some catchments. We estimated removal from off-network stream segments as the median removal of all first order streams in the same catchment. In this case we are assuming these first order segments behave similarly to other first order stream reaches. Off-network ditches and canals, identified as such in the NHD dataset, are expected to provide little N removal due to their designed intent to move water efficiently, reducing residence time (Buchanan et al. 2013), and therefore assign removal as the lower quartile of removal from the highest order streams in the same catchment. Finally, off-network lakes and ponds are assigned N removal at the 3rd quartile of removal from all lakes and ponds in the same catchment [Hampton et al. (2019); ].

## Caveats

N-Sink was designed for catchments with a surface water network. It is not intended in arid areas or where surface water flow is highly manipulated, such as in urbanized or tile-drained agricultural areas. In these situations surface water flow is highly variable and often difficult to estimate or predict.

Here we present two examples using the static maps generated by the nsink package to illustrate how we envision this tool being useful in decision-making at the local level.

# Use Examples

N-Sink is now available for HUC-12 catchments in coastal CT, RI and southern MA and decision-makers are beginning to make use of maps and data available as an [interactive tool](https://uconnclear.maps.arcgis.com/apps/MapSeries/index.html?appid=e96b01502f5b4aeda64aacf0cb5234c7&folderid=dde199a7cf674845a83b43a3011b1d31). However, we are aware of various ways this tool is starting to be integrated into risk analyses. Here we demonstrate two examples using the Niantic River (CT) the Farm River (CT) HUCs.

# Results

# Discussion and conclusions

Really great conclusions and discussion thereof goes here!

# Bibliography

Alexander, R. B., E. W. Boyer, R. A. Smith, G. E. Schwarz, and R. B. Moore. 2007. The role of headwater streams in downstream water quality 1. JAWRA Journal of the American Water Resources Association 43:41–59.

Buchanan, B., K. Falbo, R. Schneider, Z. Easton, and M. Walter. 2013. Hydrological impact of roadside ditches in an agricultural watershed in central new york: Implications for non-point source pollutant transport. Hydrological processes 27:2422–2437.

Hampton, T. B., J. P. Zarnetske, M. A. Briggs, K. Singha, J. W. Harvey, F. D. Day-Lewis, F. MahmoodPoor Dehkordy, and J. W. Lane. 2019. Residence time controls on the fate of nitrogen in flow-through lakebed sediments. Journal of Geophysical Research: Biogeosciences 124:689–707.

Howarth, R. W., and R. Marino. 2006. Nitrogen as the limiting nutrient for eutrophication in coastal marine ecosystems: Evolving views over three decades. Limnology and Oceanography 51:364–376.

Kellogg, D., A. J. Gold, S. Cox, K. Addy, and P. V. August. 2010. A geospatial approach for assessing denitrification sinks within lower-order catchments. Ecological Engineering 36:1596–1606.

Klocker, C. A., S. S. Kaushal, P. M. Groffman, P. M. Mayer, and R. P. Morgan. 2009. Nitrogen uptake and denitrification in restored and unrestored streams in urban maryland, USA. Aquatic Sciences 71:411–424.

Mayer, P. M., S. K. Reynolds Jr, M. D. McCutchen, and T. J. Canfield. 2007. Meta-analysis of nitrogen removal in riparian buffers. Journal of environmental quality 36:1172–1180.

Mengistu, S. G., H. E. Golden, C. R. Lane, J. R. Christensen, M. L. Wine, E. D’Amico, A. Prues, S. G. Leibowitz, J. E. Compton, M. H. Weber, and others. 2020. Wetland flowpaths mediate nitrogen and phosphorus concentrations across the upper mississippi river basin. JAWRA Journal of the American Water Resources Association.

Nixon, S. W. 1995. Coastal marine eutrophication: A definition, social causes, and future concerns. Ophelia 41:199–219.

Ryther, J. H., and W. M. Dunstan. 1971. Nitrogen, phosphorus, and eutrophication in the coastal marine environment. Science 171:1008–1013.

Seitzinger, S., J. A. Harrison, J. Böhlke, A. Bouwman, R. Lowrance, B. Peterson, C. Tobias, and G. V. Drecht. 2006. Denitrification across landscapes and waterscapes: A synthesis. Ecological applications 16:2064–2090.