**Rpackage: StreamNetworkTools**

Darin Kopp | [darinkopp@gmail.com](mailto:darinkopp@gmail.com)

**Purpose:** an R package to derive covariates from NHDPlusV2 dataset and facilitate continental scale analyses of river networks and national scale monitoring efforts.

**Status:** Available asprovisional source code

**Data Requirements**:

NHDPlus documentation <http://www.horizon-systems.com/NHDPlus/NHDPlusV2_documentation.php>

/NHDSnapshot/Hydrography/NHDFlowline.dbf

/NHDSnapshot/Hydrography/NHDFlowline.shp

/NHDPlusCatchment/Catchment.shp

/NHDPlusAttributes/PlusFlowlineVAA.dbf

/NHDPlusAttributes/PlusFlow.dbf

/NHDPlusAttributes/elevslope.dbf

/VPUAttributeExtension/CumTotPrecip

/VPUAttributeExtension/CumTotTemp

/VPUAttributeExtension/CumTotROMA

/VPUAttributeExtension/CumTotNLCD

/VogelExtension/vogelflow.dbf

/EROMExtension/EROM\_MA0001.DBF

**Dependencies:** devtools**,** foreign, sf, legeom, dplyr, sp

**Installing StreamNetworkTools\***

1. library("devtools")

install\_git("https://github.com/dkopp3/StreamNetworkTools\_git.git", subdir = "StreamNetworkTools")

1. install\_github(“dkopp3/StreamNetworkTools\_git”, subdir = "StreamNetworkTools")

\*if reinstalling, remember to:

1. remove.packages("StreamNetworkTools")
2. restart R

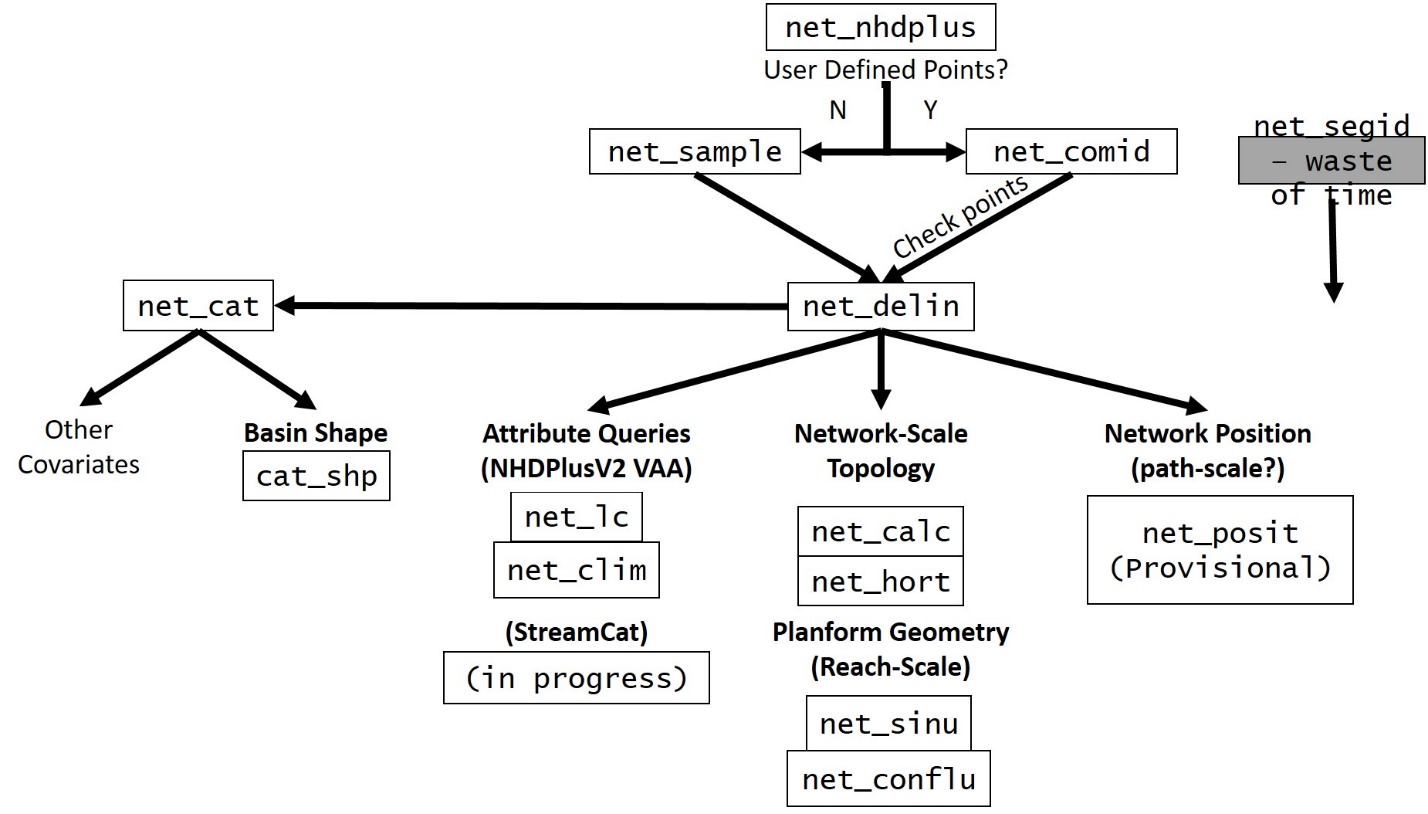


Figure 1: Workflow Schematic of StreamNetworkTools. The used begins by downloading NHDplusV2 Data for specified vector processing unit (VPU) using net\_nhdplus. Depending on whether there are specific sampling points/networks of interest, net\_sample or net\_comid can be used to identify NHDPlus COMID. Once COMIDs have been identified, the net\_delin, uses upstream navigation to delineate all COMIDs upstream from the root. The net\_delin object is required for all subsequent processing steps. net\_calc queries NHDPlus Catchments associated with a net\_delin COMID. net\_lc and net\_clim summarize landcover percentages and climatic variables for the network, respectively. net\_calc and net\_hort characterize network-scale geometry and net\_sinu and net\_conflu identify planform geometry assocaiated with reaches and confluences. net\_posit is underdevelopment but will create distance for points within the same network.

**Existing Functions:**

**Data Preparation**

net\_nhdplus(nhdplus\_path, download, vpu, files, zip\_7 = "C:/Program Files/7-Zip")

**net\_sample (nhdplus\_path, vpu, ws\_order, n)** – returns vector of comid’s from NHDPlusV2 to serve as root node of network specified by a given order.

nhdplus\_path = character string identifying path to NHDPlusV2 dataset “./ /NHDPlus”

vpu = character identifying the vector processing unit of interest

ws\_order = numeric identifying the order of the network

n = number of comids extracted form VPU

**net\_comid (pts, CRS, nhdplus\_path, vpu, maxdist)** – Identifies NHDPlusV2 flowline (COMID) within a user-defined search radius (maxdist) around a sampling point. Returns distance to nearest flowline vertex ($snap\_dist (m)), coordinates of flowline vertex ($snap\_X, $snap\_Y, (CRS=5070)), NHDPlusV2 Common Identifier ($COMID), Geographic Names Information System Names of flowline ($GNIS\_NAME) and vector processing unit ($vpu). Recommend visual inspection of points to ensure they were matched to correct flowline. The COMIDs here will serve as the root for stream networks. Ignore warning message. Points that occur further away from specified stream than “maxdist” will not appear in output.

**net\_delin (group\_comid, vpu, nhdplus\_path)** – delineates full network from root comid specified by net\_comid(…), net\_sample(…) or the user. Returns list with comid dataframe ($Network) with all comids associated with the network indexed by the root comid, nested position data frame ($Nested\_Position) identifying networks with multiple, “flow-connected” sampling points and simple feature ($SF\_Obj) with all independent networks.

**net\_cat (netdelin, vpu, nhdplus\_path, dissolve)** – extracts NHDPlusV2 catchments associated with network flowlines. Dissolve specifies whether to merge sub catchments for each network comid. Returns a list with a simple feature ($sf\_obj) of NHDPlusV2 catchments and a dataframe of comid’s without catchments ($na\_cat) - NHDPlusV2 did not delineate catchments for small flowlines (see manual – will need to check into solution). Catchments which could not be joined to flowline were not included, leaving “holes” in basin polygon when dissolve = “Y”. These holes seem to be consistent with Value Added Attribute calculations in NHDPlusV2. Lastly, for further use of these catchments in ArcGIS, run “repair geometry” in ArcGIS to avoid errors. Fixing this soon

**net\_calc (netdelin, vpu, nhdplus\_path) –** summarizes network characteristics: $Q001E (cf) is estimated mean annual discharge at root node, $V0001E (cfs) estimated mean annual velocity at root node, MAFLOWV discharge estimated by vogel method, MAVELV velocity estimated by vogel method. $maxelev (cm) highest elevation of a flowline, $minelev (cm) lowest elevation of flowline, $relief (cm) max elevation minus min elevation, $slope.wt length weighted slope of network, $slope\_NA count of reaches without slope value (-9999 or -9998), $head.h2o number of headwaters (i.e. terminal nodes), $AREASQKM.x (km^2) drainage area of network, $LENGTHKM (km) total length of network, $drain.den total length divided by drainage area, $reach.cnt total number of segments (i.e. major flow path and stream order) $diver.cnt number of divergences.

**net\_hort (nhdplus\_path, vpu, netdelin) –** calculates Horton ratios and network number, lengths and drainage area for a netdelin object. Orders are defined by unique combination of "STREAMORDE" and "LEVELPATHI" (potentially subject to change). Returns a two element list. $topology stores stream order, stream length and drainage area for each stream order extending from a root comid. $Horton\_est estimates the Horton ratios (“laws of stream numbers, lengths and areas) using simple linear models. Since a given a root COMID may occur in the middle of a reach, ohm represents the stream order used in the calculations (max(WS.ord)-1).

**net\_conflu (netdelin, nhdplus\_path) –** identifies locations of tributary junctions ($X)($Y) within a network and calculates the confluence angle ($alpha) and ratio between mainstem and tributary ($area\_ratio). In current state angles and ratios are not calculated for complex junstions ($complex)

**net\_sinu (netdelin)** – calculates mean sinuosity (total length/straight line length) of a network from netdelin object

**net\_clim (nhdplus\_path, vpu, netdelin) –** summarizes NHDPlusV2 climate data derived from PRISM 1971-2000 into BioClim-type metrics. Where applicable, temperature units are (deg C) and precipitation units are (mm). MISSDATAA.X missing area of temperature data, $TEMPVC mean annual temperature of network, $seasonality\_t, standard deviation of mean monthly temperatures, $warm\_mo month with highest mean temperature $warm\_mo\_t mean temperature of warmest month $cold\_mo month with lowest mean monthly temperature, $cold\_mo\_t mean temperature of coldest month, $diff\_t temperature difference between warmest and coldest months $warm\_q\_t mean temperature of warmest quarter (3 consecutive months with highest mean

temperature) $warm\_q warmest quarter, $cold\_q\_t mean temperature of coldest quarter (3 consecutive months with lowest mean temperature), $cold\_q coldest quarter $MISSDATAA.y missing data area for precipitation $PRECIPVC mean annual cumulative precipitation of network, $wet\_mo wettest month $wet\_mo\_p precipitation of wettest month $dry\_mo driest month $dry\_mo\_p precipitation of driest month $seasonality\_p coefficient of variation of mean monthly precipitation, $wet\_q\_p mean cumulative precipitation of wettest quarter $wet\_q wettest quarter, $dry\_q\_p precipitation of driest quarter, $dry\_q driest quarter, $dry\_q\_t mean temperature of driest quarter $wet\_q\_t mean temperature of wettest quarter, $warm\_q\_p mean cumulative precipitation of warmest quarter, $cold\_q\_p mean cumulative precipitation of coldest quarter $RUNOFFVC mean annual cumulative runoff

**net\_lc (nhdplus\_path, vpu, netdelin) –** calculates percent landcover from NLCD 2011 for catchment area

**Error Log:**

Upon loading StreamNetworkTools: “package or namespace load failed for 'dplyr' in loadNamespace…” because of missing package “bindr” resolved with restarting R and install.packages(“bindr”) before loading StreamNetworkTools

NHDPlus stream Order Error: Group id = 17153302, vpu = 18 is listed as 4th order but contains 5th order network upstream. Error at Tributaty junction of COMID’s 17153404 (STREAMORDE = 3) and 948030237 (STREAMORDE = 5) yields 3 order stream, should be 5th order. Needs to be reported