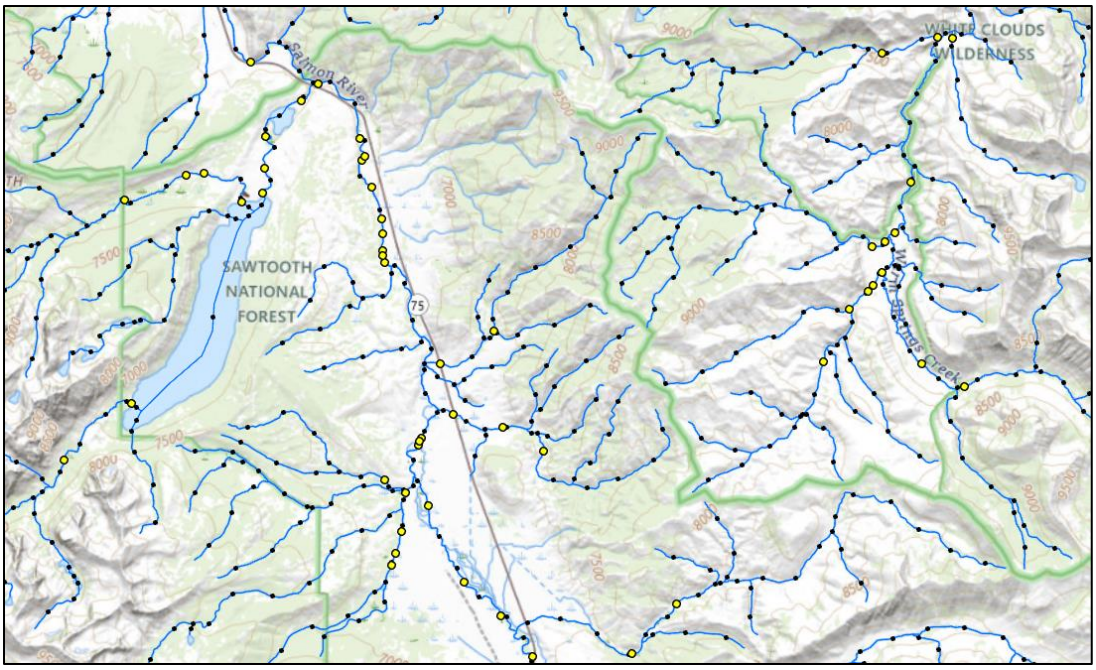
Preparing Input Data for Spatial Stream Network Modeling

David Nagel[1] and Erin Peterson[2]

U.S. Forest Service, Rocky Mountain Research Station
Water and Watersheds Program
Boise Aquatic Science Laboratory
322 E. Front St., Boise, ID

December 29, 2025



UNITED STATES DEPARTMENT OF AGRICULTURE - U.S. FOREST SERVICE
Rocky Mountain Research Station
Air, Water, & Aquatic Environments Program

Boise Spatial Streams Group

Author affiliations: [1] US Forest Service, Rocky Mountain Research Station, Water and Watersheds Program, 322 E. Front St., Suite 401, Boise, ID 83702. [2] EP Consulting, Brisbane, Australia.

Version 12-29-2025

1.0 Introduction

1.1 Overview

To characterize stream resources, physical, chemical, and biological observations are frequently sampled at discreet locations along a stream network. However, extrapolating these data to yield semi-continuous predictions throughout the network has proved challenging using standard, non-spatial statistical techniques developed for terrestrial systems. Two R Statistical Software (R Core Team 2024) packages have been developed to address these issues: SSNbler (Peterson et al. 2024) and SSN2 (Dumelle et al. 2023). The SSNbler package is used to generate and format the data prior to modeling in R, while the SSN2 package is used to fit the models and generate predictions at unsampled locations with associated estimates of uncertainty.

The SSNbler package requires a vector stream network in a standard format such as Esri shapefile or GeoPackage. The package also requires a point dataset representing an observed variable (response variable), generally collected from field measurements, representing themes such as pollutant concentrations, stream temperature, or aquatic species density. One or more prediction point datasets are useful for extrapolating predicted responses throughout the network, but this component is optional. These three components (stream network, observation points, and prediction points) make up the core spatial components of the input data for the SSNbler package and SSN2 modeling.

To explain variance in the observation dataset and generate predictions throughout the stream network, a set of covariates or predictors (independent variables) is necessary. These covariates may include spatial variables such as elevation, stream discharge, or land cover to help explain variation in the observation dataset. These covariate values must be assigned to both the observation and prediction point datasets in the feature attribute tables.

SSNbler ingests the streamline network, observation values, and associated covariate values, generates spatial measures needed for modelling, and outputs the SSN object used in the SSN2 package. The SSN2 package is then used to fit spatial stream-network models to the data and generate predictions of the response variable at the prediction points throughout the network.

1.2 Purpose

The purpose of this document is to describe a recommended framework for formatting the spatial datasets required by the SSNbler and SSN2 packages. This framework includes structuring the observation and prediction point features, building an attribute table ID structure, and guiding the covariate attribution process. It is important to note that the stream network must be topologically sound before this process can begin. One should also note that this document only describes a

suggested framework for organizing the input data. Because the software tools are flexible, other organizational structures may be more appropriate for individual applications.

2.0 Software Requirements

A full featured GIS software suite is required to process the spatial data and attribute tables required for the SSNbler package. ArcGIS Pro and QGIS are two packages recommended by the authors. Other packages may be equally sufficient.

3.0 Data Overview

3.1 Nomenclature

The SSNbler and SSN2 packages enable users to fit statistically valid models of aquatic phenomenon on stream networks and make subsequent predictions of aquatic attributes at unsampled locations. Field-collected data are needed to fit the spatial stream-network model and these data are referred to as *observed data*, which are collected at *observation sites* (Figure 1). When a statistical model is fit to an observed variable, we refer to it as the *response variable* (i.e., dependent variable). In most cases, it is also desirable to include *covariates* in the model (i.e., predictors or independent variables), which help explain variability in the response. For example, if the response variable is stream temperature, then canopy cover might be used as a covariate in the model. A fitted model can also be used to generate predictions at user-defined points along the stream network. We refer to these as *prediction points*.

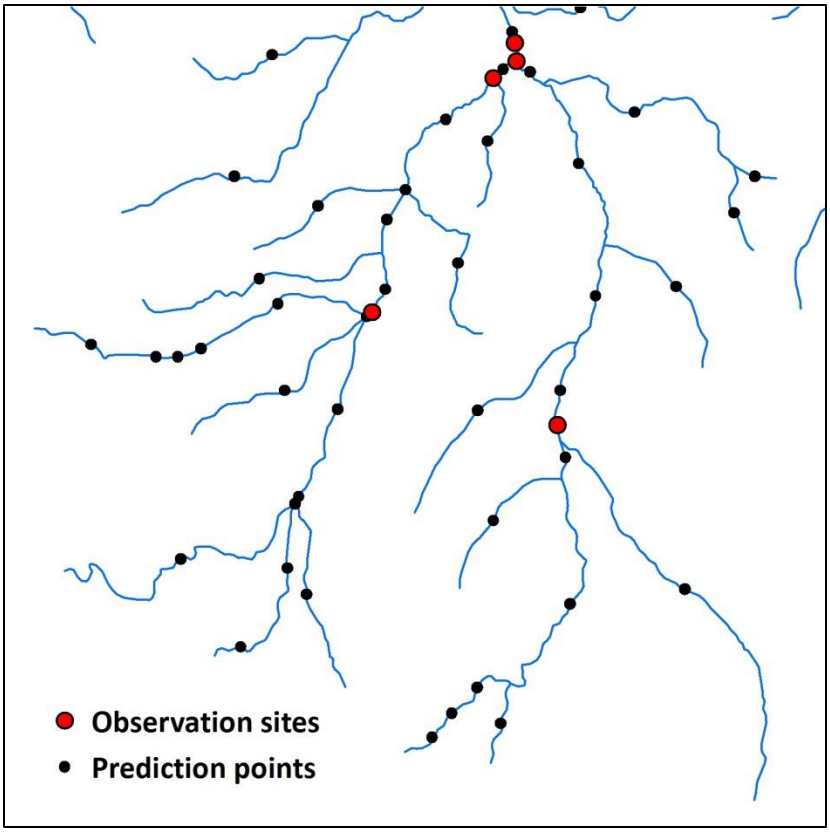


Figure 1. Example observation sites and prediction points. These prediction points are assigned at the midpoint of each stream segment. Thus, each stream segment can be assigned a response variable.

3.2 Input Data

3.2.1 Stream Network Data

Stream network data may come from a variety of sources. A popular dataset in the United States is the USGS NHDPlus High Resolution (HR) dataset. These 1:24,000 scale streamlines from the original topographic quadrangle maps are attributed with a variety of value-added attributes that may be useful as covariates in the SSN2 modeling process. The data are not topologically corrected however and require editing.

The USGS NHDPlusV2 dataset is a digital representation of the 1:100,000 scale map streamlines and like NHDPlus HR, includes value-added attributes that may be useful for modeling purposes. This dataset is associated with an even larger set of stream attributes produced by EPA, called the StreamCat dataset, substantially extending the utility of NHDPlusV2 for modeling purposes. The StreamCat data may be accessed here: https://www.epa.gov/national-aquatic-resource-surveys/streamcat-dataset

The NHDPlusV2 dataset was modified by the Forest Service, RMRS, Boise Aquatic Sciences Lab, for use with the SSNbler and SSN2 packages. This dataset is called the National Stream Internet (NSI). The NSI dataset consists of streamlines and prediction points. The NHD digital streamlines, or *NHDFlowline* data, were edited to conform to topological constraints required by the SSNbler and SSN2 tools. The lines alone are referred to as the NSI network. In addition, prediction points have been generated at the midpoint of each NSI streamline that contains a unique ComID value. The data are available for the conterminous U.S. The data may be freely downloaded from here: https://research.fs.usda.gov/rmrs/projects/national-stream-internet#overview

Streamline datasets may also be derived from digital elevation models or LiDAR data, such as streamlines from the USGS 3D Hydrography Program. These high-resolution data are not widely available at this time, but the coverage is expanding and should be completed for the conterminous US within the next decade. The streams dataset must be in a rectangular coordinate system such as UTM or Albers, rather than a geographic coordinate system such as decimal degrees (i.e., latitude and longitude coordinate systems are not allowed).

3.2.2 Observation Data

Spatially referenced observed data are required to generate predictions using the SSN2 package. Observed data generally include field-collected measurements such as temperature, chemistry, fish presence/absence or species count data, or any other variable that can be estimated using one or more covariates. The observation sites are represented as GIS point locations and must be in the same map projection as the stream network. These data may be acquired from various resource agencies or obtained through field data collection efforts.

3.2.3 Covariate Data

Covariate data are used to explain variability in the response variable during the modeling process. For example, if stream temperature is the response variable, suitable covariates may include elevation, canopy cover, and drainage area. Covariate data must be available for every observation measurement in the dataset. If the SSN2 package is used to make predictions, then covariates must also be available at each prediction site and stored in the attribute table.

The covariates are typically derived from GIS layers that are available throughout the entire stream network. Alternatively, attributed stream datasets, such as NHDPlusV2 and StreamCat, contain many such variables that may be joined to the NSI streamline dataset.

3.2.4 Prediction Points

Prediction points are user defined locations along the stream network where the response variable can be estimated (Figure 1). If a spatial stream-network model is used to make predictions, these predictions can be linked in a GIS to the prediction points and also the streams data through the 'rid' field created in SSNbler.

It's important to carefully consider a prediction point design strategy that represents the underlying data being modeled. For example, prediction points may be generated at an equal interval throughout the network, such as 100m or 1 km, or points may be generated at the midpoint of each stream reach, yielding a 1:1 relationship between prediction points and stream reaches. Increasing the number and spatial density of prediction sites increases pre-processing time in SSNbler. Therefore, the spatial density of the prediction points should reflect the spatial resolution of the covariate data and also the purpose of the spatial stream-network model. For example, it may not be worthwhile to generate prediction sites at 100m intervals if the covariates are only available at the line feature scale. In that case, prediction points at the mid-point of each stream segment may be sufficient. However, prediction sites at 100m intervals may be appropriate in some portions of the network if the aim is to generate abundance estimates using block kriging (e.g. Isaak et al. 2017).

As mentioned previously, the NSI dataset includes a set of prediction points at the midpoint of each streamline segment. There are also several GIS tools that can be used to generate prediction points at specific intervals in Esri ArcGIS Pro (e.g., Points Along Line tool, https://pro.arcgis.com/en/pro-app/latest/help/editing/create-point-features-along-a-line.htm) and QGIS (native pointsalonglines tool or gdal pointsalonglines tool).

3.2.4 All Spatial Data. Important.

It is critical that all streamlines and points are in the same map projection. In addition, the projection must in a rectangular coordinate system such as UTM, state plane, or Albers. A geographic coordinate system such as decimal degrees will not work with the SNNbler and SSN2 packages because distances cannot be computed accurately when generating the network metrics.

4.0 Preparing the Spatial Data

Example Study Area

For illustration purposes, we'll be looking at a small drainage called Fourth of July Creek in the Salmon River Basin, Idaho. The drainage length is approximately 20 km, east to west. Observation sites are yellow and prediction points are black. There are 11 observation sites with a total of 35 stream

temperature records. There are 50 prediction points. Prediction points are spaced at an interval of approximately 1 km. The streamline data have already been topologically corrected (Figure 2).

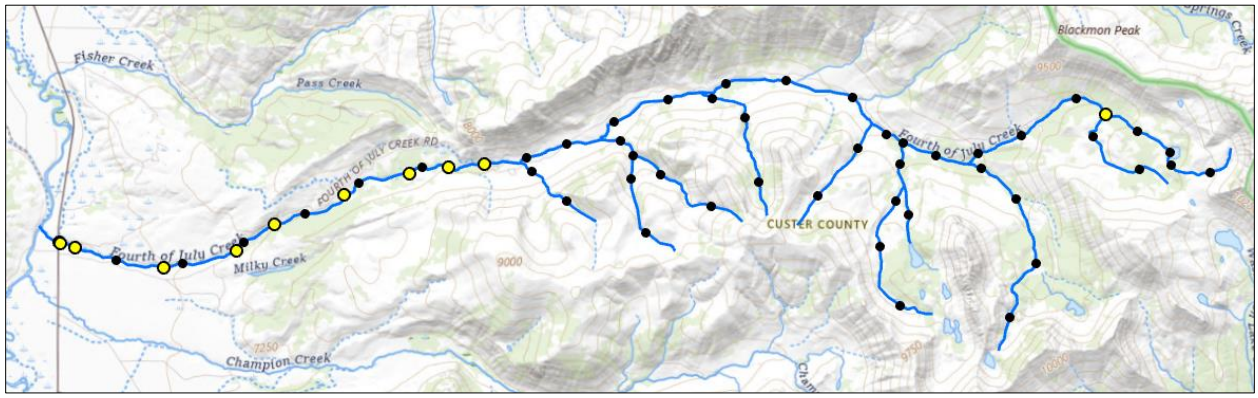
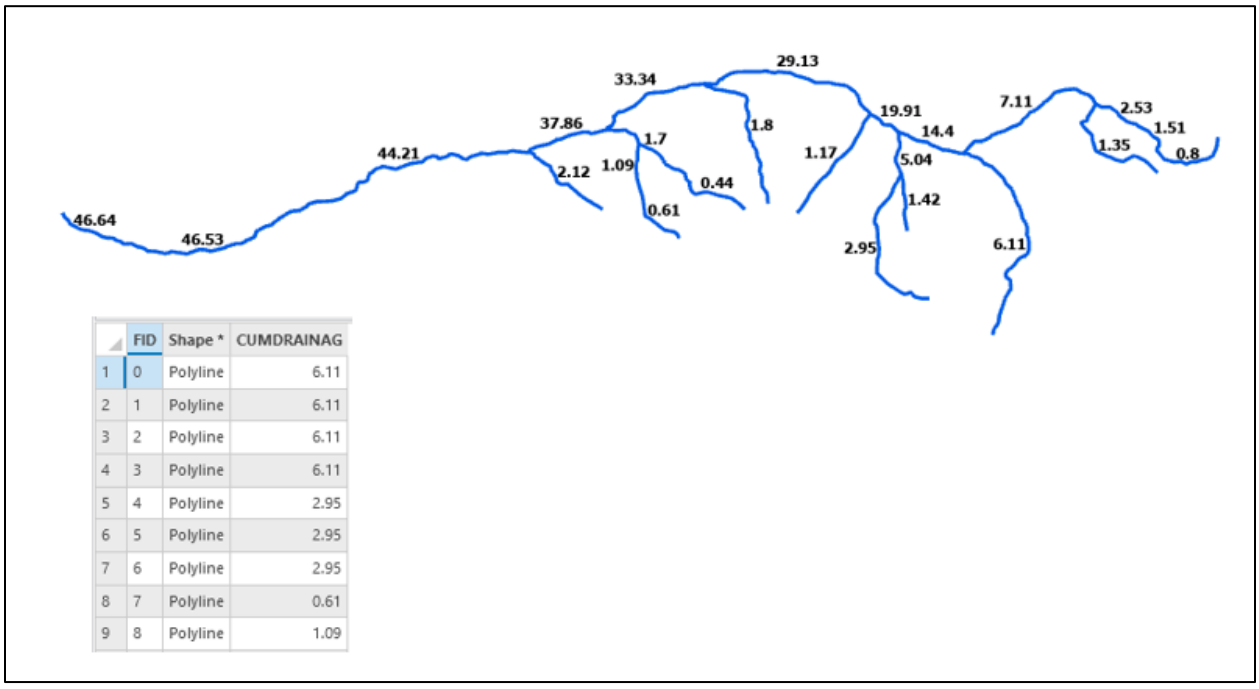


Figure 2. Example study site. Yellow points represent observation sites and black points represent prediction points.

4.1 Attribute Streamlines with Weighting Variable

A weighting variable is required to generate the SSN. This variable assigns a relative weight to each stream reach, ensuring that, for example, where a small tributary intersects a larger mainstem reach, the mainstem imparts greater influence in the statistical model.

Cumulative drainage area is often used as a common weighting variable. In the NHDPlus dataset, this attribute is called TotDASqKM (total drainage area, square kilometers). In this example, each stream reach is assigned the weighting variable, CUMDRAINAG (cumulative drainage area).



	FID	Shape *	CUMDRAINAG
1	0	Polyline	6.11
2	1	Polyline	6.11
3	2	Polyline	6.11
4	3	Polyline	6.11
5	4	Polyline	2.95
6	5	Polyline	2.95
7	6	Polyline	2.95
8	7	Polyline	0.61
9	8	Polyline	1.09

4.2 Observation Data

Observation data are a core component of the spatial statistical modeling process. In this example, we'll use stream temperature as our response variable. In other words, we'll utilize stream temperature sample data from field measurements at a limited number of discreet locations, and then use the modeling tools to extrapolate stream temperature to the rest of the network at the prediction points. Note that we won't model stream temperature here, but only illustrate the data preparation process.

The first step is to collect the necessary field data and be sure that those locations are spatially referenced. The locations should fall on, or very close to, the digital streamlines. The SSNbler package will snap observation and prediction points to the streamlines, however it's prudent to snap or manually move points close to the streamlines ahead of time to be sure all points get snapped to the correct streams. In Figure 3 below, the observation points have been snapped to the digital streamlines, even though the streamlines are not perfectly coincident with the stream channel location in the field.

Figure 3. Observation sites snapped to the digital streamline.

A stream temperature record must be associated with each observation point. In Figure 4 below is an example of the attribute table for the observation sites.

	FID	Shape *	OBSPRED_ID	SITE_ID	SampleYear	MeanAugT
1	0	Point	1	303	2001	9.107419
2	1	Point	2	303	2002	12.709355
3	2	Point	3	303	2003	8.77129
4	3	Point	4	303	2004	8.351724
5	4	Point	5	303	2006	8.603793
6	5	Point	6	303	2007	8.955161
7	7	Point	8	303	2012	8.197742
8	8	Point	9	303	2014	8.603548
9	9	Point	10	303	2020	9.592667
10	10	Point	11	9402	1994	11.782903
11	11	Point	12	9402	2001	14.899032
12	12	Point	13	9402	2002	9.290323
13	13	Point	14	9402	2004	9.830667
14	14	Point	15	9402	2014	9.257419

Figure 4. Observation site attribute table. FourthOfJulyCreek_ObservationSites.shp

Looking at the attribute table fields, notice that the first field is named OBSPRED_ID. This field contains a unique ID number for each temperature record. The field is called OBSPRED because it will eventually hold a unique ID for both the observation points and prediction points, once the two feature classes are merged. We want to be sure that no observation points or prediction points have the same ID.

The next field is called SITE_ID. In Figure 5 below, the two observation sites highlighted in magenta represent sites with SITE_ID 303 and 9402 shown in Figure 4. Note that both sites have multiple temperature records associated with them, collected in different years, and recorded in the SampleYear field.

Figure 5. SITE_IDs 303 and 9402 highlighted in magenta.

Finally, the MeanAugT field records the field collected mean August stream temperature generated from observed temperature records collected at each site, for each year. Note that SITE_ID = 303 has nine temperature records. In the spatial dataset, this location also has nine point features. However, the points are all at the exact same location and intersect each other, making them appear as a single point.

This is the basic structure for the observation site geometry and attribute table. There is a point feature associated with each temperature record, and records at the same location receive the same SITE_ID. *Important note: If your data does not have multiple observations at a single location, the SITE_ID field is not strictly necessary.*

4.2 Prediction Points

Prediction points allow the user to extrapolate the observation data throughout the network after a model is developed with the SSN2 package. Developing a prediction point strategy is an important step. In the example here, prediction points were generated at approximate 1 km intervals along the stream network. In addition, the stream network was divided into 1 km reaches, creating a 1:1 relationship between each prediction point and stream reach. This 1:1 relationship is not required to make predictions at the points, however this design is useful for easily mapping predictions back to the stream reaches if desired. Figure 6 below illustrates that each prediction point is associated with a matching stream reach. Prediction points can also be placed at each stream reach midpoint, or using another criteria chosen by the user.

As mentioned previously, the NSI dataset includes a set of prediction points centered at the reach midpoints. There are also several GIS tools that can be used to generate prediction points at specific intervals in Esri ArcGIS Pro (e.g., Points Along Line tool, https://pro.arcgis.com/en/pro-app/latest/help/editing/create-point-features-along-a-line.htm) and QGIS (native pointsalonglines tool or gdal pointsalonglines tool).

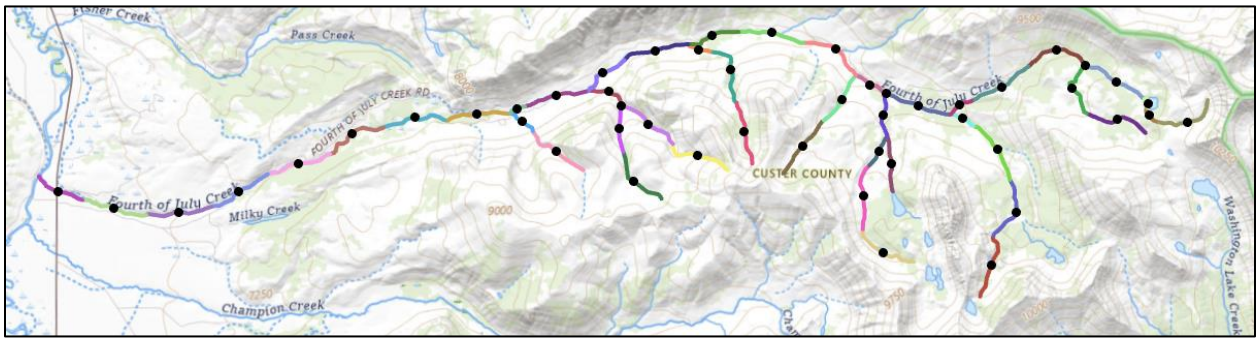


Figure 6. Each prediction point with a matching stream reach.

There are 50 prediction points in this example. Note in Figure 7 below that there is only one field in the attribute table, OBSPRED_ID. Notice also that OBSPRED_ID is the first field found in the observation point feature class (Figure 4). The records for OBSPRED_ID are populated sequentially, starting with value 100,000. This number system ensures that observation and prediction point IDs will be unique, assuming there are fewer than 100,000 observation records. OBSPRED_ID can be computed by calculating the field to FID + 100000.

FID	Shape *	OBSPRED_ID
0	Point	100000
1	Point	100001
2	Point	100002
3	Point	100003
4	Point	100004
5	Point	100005
6	Point	100006
7	Point	100007
8	Point	100008
9	Point	100009
10	Point	100010
11	Point	100011
12	Point	100012

Figure 7. Prediction points attribute table. FourthOfJulyCreek_PredictionPoints.shp

4.3 Covariates

Covariates are additional data layers that help explain variation in the response variable and are used to make predictions throughout the stream network. For example, an important covariate for predicting stream temperature is elevation. Where elevations are high, stream temperatures tend to be cooler. Both the observation site and prediction points will be attributed with the covariate data.

Covariates must be available for the entire stream network being modeled. Sources of covariate data are beyond the scope of this document, but numerous datasets are available nationally for the US that could serve this purpose.

4.3.1 Merge observation and prediction feature classes

Since covariates must be attributed to both the observation sites and prediction points, it's expedient to merge the two spatial datasets and assign the covariates to the single merged dataset. Merging the observation and prediction points at this juncture also ensures that the two datasets have the exact same field names in the exact same order. The two must have identical field names and order to work with the SSNbler and SSN2 packages.

When doing the merger, the observation points should be entered first, so that they appear at the top of the attribute table in the merged feature class.

Note Figure 8 below showing the merged observation and prediction dataset (hereafter 'all-sites'). The table shows the transition point between observed points and prediction points. Notice that the OBSPRED_ID changes from 35 to 100000 at this point. Also note that the field SITE_ID, SampleYear, and MeanAugT are Null for the prediction points, as they should be.

	OBJECTID *	Shape *	OBSPRED_ID	SITE_ID	SampleYear	MeanAugT
27	27	Point	27	11191	2015	9.525484
28	28	Point	28	11548	2000	11.640968
29	29	Point	29	39218	2014	8.787097
30	30	Point	30	39244	2009	7.94871
31	31	Point	31	39244	2014	7.955806
32	32	Point	32	39245	2014	8.316129
33	33	Point	33	39261	2015	10.104516
34	34	Point	34	39277	2013	8.899032
35	35	Point	35	39277	2014	8.546129
36	36	Point	100000	<Null>	<Null>	<Null>
37	37	Point	100001	<Null>	<Null>	<Null>
38	38	Point	100002	<Null>	<Null>	<Null>
39	39	Point	100003	<Null>	<Null>	<Null>
40	40	Point	100004	<Null>	<Null>	<Null>
41	41	Point	100005	<Null>	<Null>	<Null>
42	42	Point	100006	<Null>	<Null>	<Null>

Figure 8. All-sites observation and prediction feature classes.

4.3.2 Assign covariates

The next step is to begin assigning covariates to the all-sites dataset. Watershed area (i.e. cumulative drainage area) is often used as a covariate in spatial stream-network models because it acts as a crude surrogate for stream flow volume, which is generally more difficult to obtain. So cumulative drainage area will be the first covariate we assign to all sites. Note that watershed area is also commonly used to generate the spatial weights used in the tail-up covariance function; although technically any numeric variable could be used as long as it is available for every line feature in the streams dataset. Variables

used to generate spatial weights must be present in the *streams* attribute table when input data are imported into SSNbler.

In Figure 9 below, cumulative drainage area was added to the all-sites dataset using a spatial join with the NHDPlusV2 dataset. The field is called CUMDRAINAG, with units in km^2. Notice that CUMDRAINAG has been assigned to both the observation and prediction points in all-sites.

	FID	Shape *	OBSPRED_ID	SITE_ID	SampleYear	MeanAugT	CUMDRAINAG
28	27	Point	28	11548	2000	11.640968	46.64
29	28	Point	29	39218	2014	8.787097	46.53
30	29	Point	30	39244	2009	7.94871	44.21
31	30	Point	31	39244	2014	7.955806	44.21
32	31	Point	32	39245	2014	8.316129	44.21
33	32	Point	33	39261	2015	10.104516	2.53
34	33	Point	34	39277	2013	8.899032	44.21
35	34	Point	35	39277	2014	8.546129	44.21
36	35	Point	100000	0	0	0	6.11
37	36	Point	100001	0	0	0	6.11
38	37	Point	100002	0	0	0	6.11
39	38	Point	100003	0	0	0	6.11
40	39	Point	100004	0	0	0	2.95
41	40	Point	100005	0	0	0	2.95
42	41	Point	100006	0	0	0	2.95

Figure 9. Adding the covariate CUMDRAINAG to all-sites, which is a merged dataset of observations and prediction sites.

Any other covariates can now be added to the dataset using a spatial join. Note in Figure 10 below that we've added elevation (ELEV), canopy cover (CANOPY) and stream gradient (SLOPE) as three new covariates for modeling stream temperature. We'll consider these the full set of covariates for this example.

Field: Add Calculate | Selection: Select By Attributes Zoom To Switch Clear De

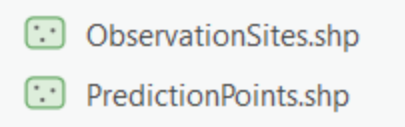
	FID	Shape *	OBSPRED_ID	SITE_ID	SampleYear	MeanAugT	CUMDRAINAG	ELEV	CANOPY	SLOPE
26	25	Point	26	11191	2005	8.333226	44.21	2261.97	32.74	0.025503
27	26	Point	27	11191	2015	9.525484	44.21	2261.97	32.74	0.025503
28	27	Point	28	11548	2000	11.640968	46.64	2071.75	0.71	0.029728
29	28	Point	29	39218	2014	8.787097	46.53	2128	5	0.03441
30	29	Point	30	39244	2009	7.94871	44.21	2261.97	32.74	0.025503
31	30	Point	31	39244	2014	7.955806	44.21	2261.97	32.74	0.025503
32	31	Point	32	39245	2014	8.316129	44.21	2232.36	24.52	0.025503
33	32	Point	33	39261	2015	10.104516	2.53	2750.7	22.33	0.098492
34	33	Point	34	39277	2013	8.899032	44.21	2203.84	29.97	0.025503
35	34	Point	35	39277	2014	8.546129	44.21	2203.84	29.97	0.025503
36	35	Point	100000	0	0	0	6.11	2797.46	12.11	0.085711
37	36	Point	100001	0	0	0	6.11	2716.91	26.9	0.085711
38	37	Point	100002	0	0	0	6.11	2620.1	24.38	0.085711
39	38	Point	100003	0	0	0	6.11	2575.64	7.13	0.085711
40	39	Point	100004	0	0	0	2.95	2719.33	6.18	0.061901
41	40	Point	100005	0	0	0	2.95	2687.67	2.23	0.061901
42	41	Point	100006	0	0	0	2.95	2623.02	2.04	0.061901

Figure 10. All-sites final covariate dataset. FourthOfJulyCreek_AllSites.shp

4.3 Preparing for SSNbler

4.3.1 Separate observations and predictions

The next step to prepare the data for the SSNbler workflow will be to split the observation sites and prediction points back into separate feature classes. To do so, simply select all OBSPRED_ID values less than 100,000 and export to a new dataset called ObservationSites.shp, or similar. Likewise select all records greater than or equal to 100000 and export to a new dataset called PredictionPoints.shp, or similar, like Figure 11 below.



ObservationSites

Field: Add Calculate Selection: Select By Attributes Zoom To Switch Clear

	FID	Shape *	OBSPRED_ID	SITE_ID	SampleYear	MeanAugT	CUMDRAINAG	ELEV	CANOPY	SLOPE
1	0	Point	1	303	2001	9.107419	44.21	2283.1	21.97	0.025503
2	1	Point	2	303	2002	12.709355	44.21	2283.1	21.97	0.025503
3	2	Point	3	303	2003	8.77129	44.21	2283.1	21.97	0.025503
4	3	Point	4	303	2004	8.351724	44.21	2283.1	21.97	0.025503
5	4	Point	5	303	2006	8.603793	44.21	2283.1	21.97	0.025503

PredictionPoints

Field: Add Calculate Selection: Select By Attributes Zoom To Switch Clear

	FID	Shape *	OBSPRED_ID	SITE_ID	SampleYear	MeanAugT	CUMDRAINAG	ELEV	CANOPY	SLOPE
1	0	Point	100000	0	0	0	6.11	2797.46	12.11	0.085711
2	1	Point	100001	0	0	0	6.11	2716.91	26.9	0.085711
3	2	Point	100002	0	0	0	6.11	2620.1	24.38	0.085711
4	3	Point	100003	0	0	0	6.11	2575.64	7.13	0.085711
5	4	Point	100004	0	0	0	2.95	2719.33	6.18	0.061901
6	5	Point	100005	0	0	0	2.95	2687.67	2.23	0.061901

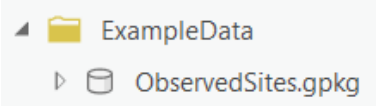
Figure 11. Observation sites and prediction points separated after adding covariates.

4.3.2. Option 1. Convert shapefile to GeoPackage format prior to using SSNbler

The observation and prediction points can be imported into the R environment using SSNbler, in either shapefile or GeoPackage format, or any other GIS format accepted by the sf R package. However, to maintain consistency with the SSNbler vignette document (An Introduction to SSNbler: Assembling Spatial Stream Network (SSN) Objects in R), guidance is provided here for converting the shapefiles to GeoPackage format.

To generate a GeoPackage from shapefile format using ArcGIS Pro:

In Catalog, right click your working folder and select New > GeoPackage. A new GeoPackage will be created.



Next in Catalog, right click the .gpkg container and select Import > Feature Classes.
Import the observation sites shapefile in the dialog box that appears.
Follow the same procedure for the prediction points and streamline shapefiles. You should have three GeoPackages, one each for observations, predictions, and streamlines, when you are finished.

To generate a GeoPackage from shapefile format using QGIS:

Add the shapefile to the QGIS viewer. Right click the layer name in the Layer pane. Select Export > Save Features As. The default is GeoPackage. Complete the dialog box and click OK. Generate one GeoPackage for each feature class, observations, predictions, and streamlines.

4.3.3. Option 2: Import a shapefile into the R environment directly using SSNbler

Do not covert to GeoPackage format, instead import shapefile format directly into the R environment.

Syntax for importing a shapefile in the R environment would look like:
ObservationSites <- st_read(paste0(path, "ObservationSites.shp")

See the document: An Introduction to SSNbler: Assembling Spatial Stream Network (SSN) Objects in R.

5.0 References

Dumelle M., Peterson E.E., Ver Hoef J.M., Pearse A. and Isaak D. (2023) SSN2: Spatial Modeling on Stream Networks in R. R package version 0.1.0. https://usepa.github.io/SSN2/

Isaak D., Ver Hoef J.M., Peterson E.E., Horan D., and Nagel D. (2017) Scalable population estimates using spatial stream-network (SSN) models, fish density surveys, and national geospatial database frameworks for streams. Canadian Journal of Fisheries and Aquatic Sciences, 74(2): 147-156.

Peterson E. E., M. Dumelle, A. R. Pearse, D. Teleki, and J. M. Ver Hoof. (2024) An Introduction to SSNbler: Assembling Spatial Stream Network (SSN) Objects in R.
https://pet221.github.io/SSNbler/articles/introduction.html

R Core Team (2021-2024) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org/. [Accessed: 07/02/2024]