

Introduction to SSN Package

Zhijiang (Van) Liu

Department of Statistics and Actuarial Sciences
University of Iowa

April 12, 2019

Acknowledgments

This presentation is summarized from the materials provided by 2018 SSN Modeling Workshop.

- Jay M. Ver Hoef
NOAA Fisheries Alaska Fisheries Science Center
- Erin E. Peterson
Queensland University of Technology
- Daniel J. Isaak
US Forest Service Rocky Mountain Research Station

Acknowledgments



Agenda

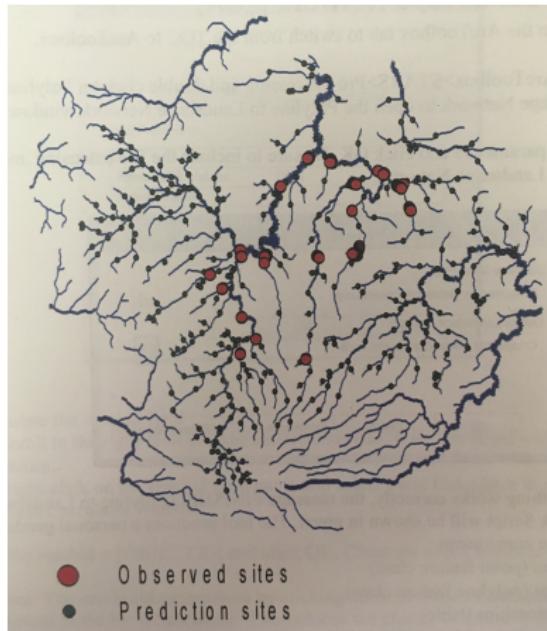
- ① Stream Network Object in ArcGIS
- ② SSN Package in R
- ③ Example: Middle Fork Data

Stream Network Object in ArcGIS

- ArcGIS is a geographic information system (GIS) for working with maps and geographic information. Starting from ArcGIS 9, it includes a *geoprocessing* environment that allows execution from any scripting language (Python is the most popular one) that supports Component Object Model standards. —— Wikipedia
- The Spatial Tools for the Analysis of River Systems (STARS) geoprocessing toolbox, created by Peterson and Ver Hoef, can be used to analyze SSN datasets.
- The data structure used in SSN package in R is inherited from the STARS tools in ArcGIS.

Stream Network Object in ArcGIS

Visualization of an example stream network dataset in ArcGIS. The observed sites are locations where data were collected, while prediction sites represent unobserved locations where predictions will be made.



Stream Network Object in ArcGIS

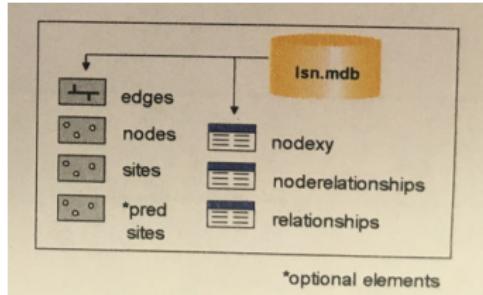
Landscape Network (LSN) is the data structure utilized by the STARS tools, its 5 essential elements are

① Nodes

- Breaks in the stream network (source, outlet, junctions, etc.)
- Each assigned a **point ID** (*pid* field)

② Edges

- Flow paths from node to node
- Each assigned a **reach ID** (*rid* field)



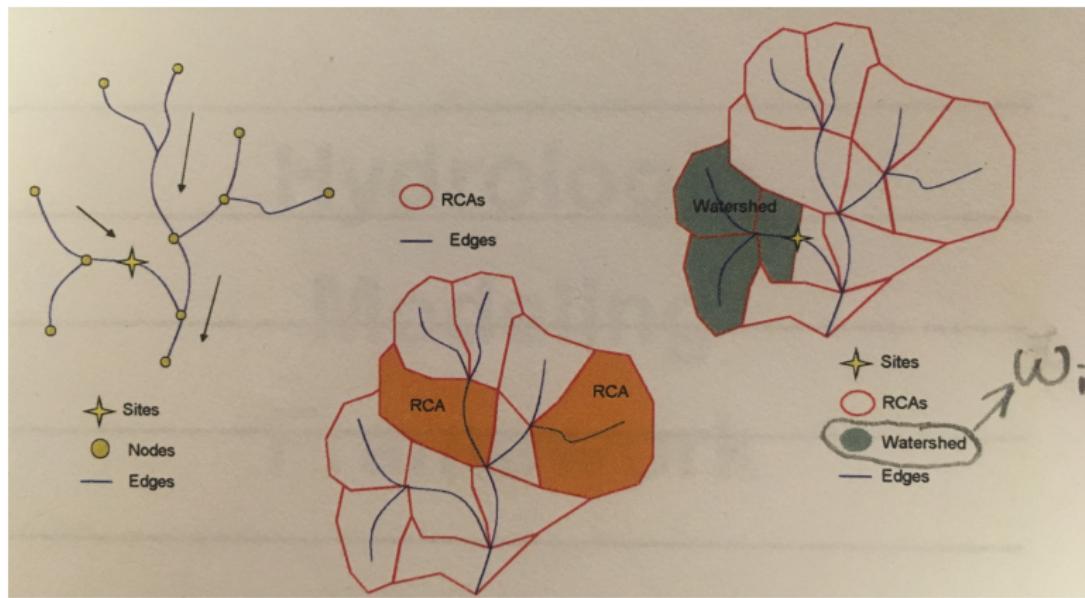
Stream Network Object in ArcGIS

- ③ nodexy table
 - Provides x,y coordinates for each node
- ④ noderelationship table
 - Provides nodes that each edge bridges (from and to), based on *rid* and *pid*
- ⑤ relationship table
 - Links edges to one another in a downstream order, based on *rid*

Additional elements are Observed sites and Prediction sites.

Stream Network Object in ArcGIS

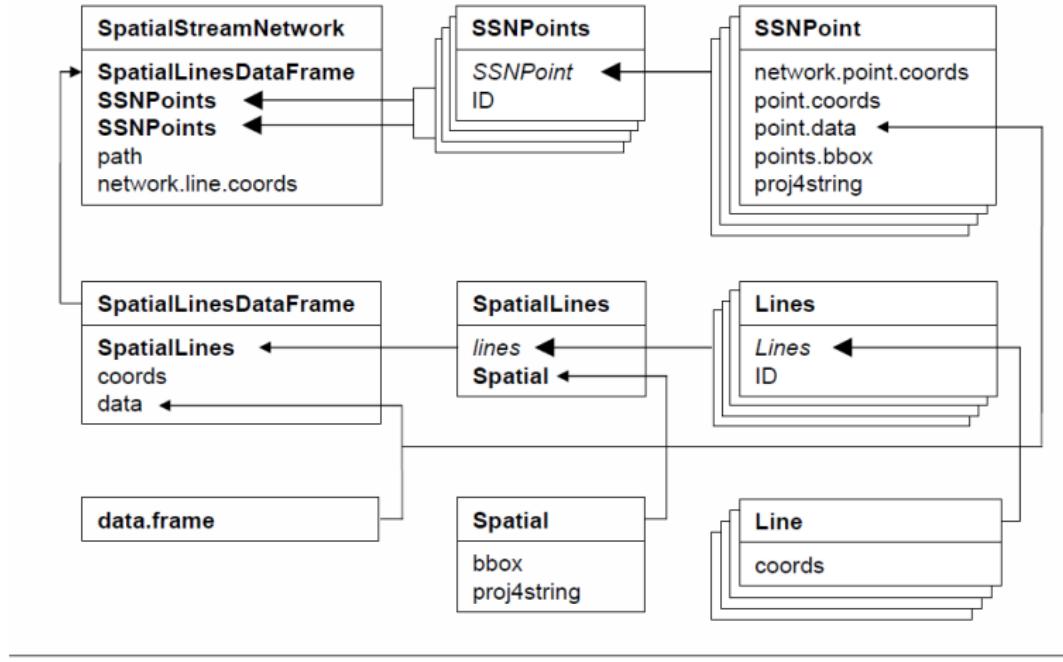
Reach Contribution Areas (RCAs) are edge-specific regions. RCAs are used to efficiently calculate watershed attributes, which is related to spatial weights (π_{ij}) in tail-up model.



SSN Package in R

- SSN package is available on CRAN, latest version is 1.1.13.
- It currently depends on R version 3.4.0 or later.
- SSN package can be installed via `install.packages` function in R or through downloaded `tar.gz` file.

SSN Package in R

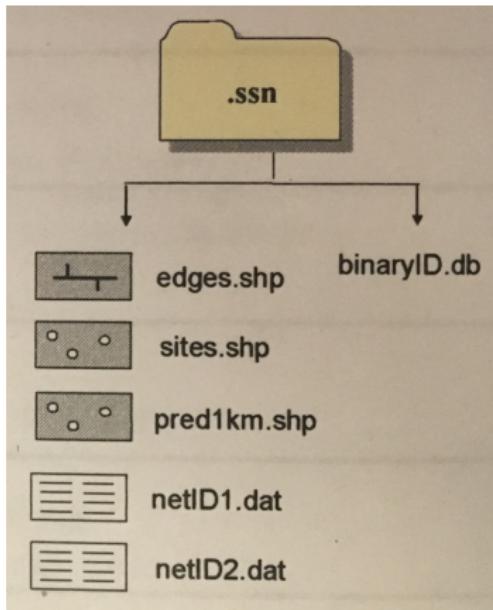


SSN Package in R

`SpatialStreamNetwork` Class is an S4 object in R.

importSSN function

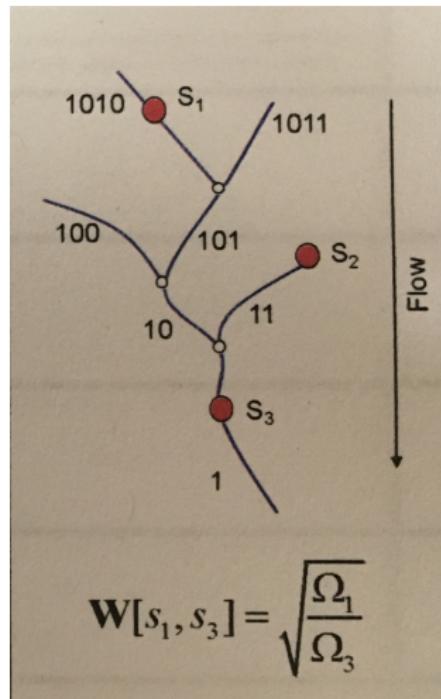
- takes SSN data stored in a “.ssn” directory and creates a `SpatialStreamNetwork` object.
- netID.dat files used to create `binaryID.db`



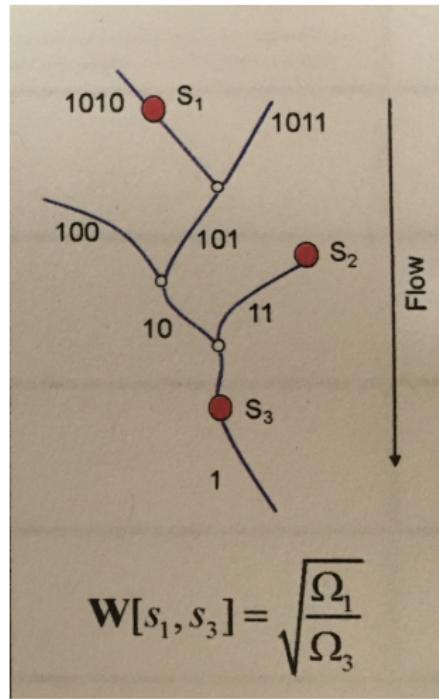
SSN Package in R

Creating Distance Matrices

- Compare network IDs
- Compare binaryIDs
 - Flow-connected: nested binary IDs
 - Flow-unconnected: binary ID diverges
- Calculate stream distance measures using upstream distances for edges & sites



SSN Package in R



Asymmetric Distance Matrix Design
 $D = \{d_{ij}\}$

- FC pairs: $d_{ij} = \text{stream distance between sites } i \text{ and } j, d_{ji} = 0.$
- FU pairs: d_{ij} and d_{ji} equal to the distances of sites i and j to their common junction.

Eg. For s_1, s_2 and s_3 in the plot,

$$D = \begin{bmatrix} 0 & 1 & 3 \\ 2 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

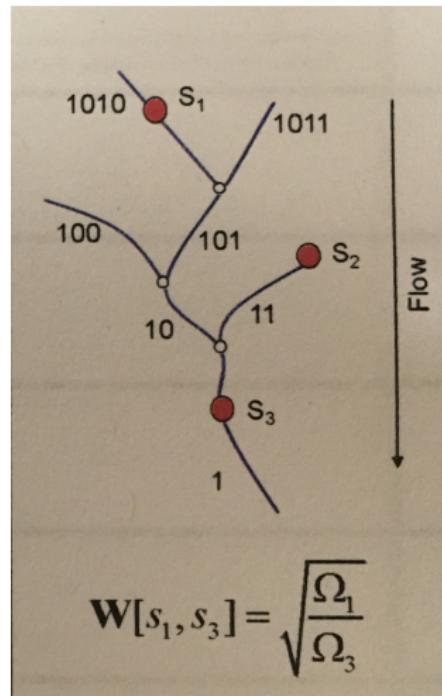
SSN Package in R

Spatial Weights

- Use binary IDs to identify FC pairs.
- Additive function values (AFVs, such as watershed areas) assigned to sites to calculate spatial weights as

$$w[s_i, s_j] = \sqrt{\frac{\Omega_i}{\Omega_j}}$$

where Ω_i is the AFV of the upstream site and Ω_j is of the downstream site.



SSN Package in R

`glmssn` object is an S3 object containing all of the information used to fit the model

- Model arguments
- Full `SpatialStreamNetwork` object
- Sample information
- Estimates
 - Standardized variables and fixed effects
 - Covariance parameters
 - Maximum likelihood estimates
 - Covariance matrix and inverse covariance matrix
- `help("glmssn-class")`

SSN Package in R

`influenceSSN` object stores influence diagnostics

- Created using `residuals()`
- Same structure as `glmssn` except
 - Raw residuals, studentized residuals, cross-validation residuals, Leverage, Cook's D, etc.
- `help("influenceSSN-class")`

`glmssn.predict` object stores predictions and standard errors

- Created using `predict()`
- Same structure as `glmssn` except
 - Predicted response variable and standard errors for the predicted sites
- `help("influenceSSN-class")`

Example: Middle Fork Data

The data set is from the Middle Fork, a stream in Idaho, United States, and is a subset of a larger data set than can be freely accessed at <http://www.fs.fed.us/rm/boise/AWAE/projects/SpatialStreamNetworks.shtml>. The data set consists of two stream networks with 45 total observations, 220 prediction locations on a 1 km spacing, 1273 prediction locations densely packed on a single stream called the Knapp, and another 654 prediction locations densely packed on a single stream called CapeHorn.

Sources

- Slides and R code of this presentation

<https://github.com/ZhijiangVanLiu/Introduction-to-SSN-Package>

- SSN package & STARS

<https://www.fs.fed.us/rm/boise/AWAE/projects/SpatialStreamNetworks.shtml>

- SSN package CRAN page

<https://cran.r-project.org/web/packages/SSN/index.html>

- SSN package tutorial

<https://www.jstatsoft.org/article/view/v056i03/v56i03.pdf>

Introduction to SSN Package

Thank you! Any Questions?

