A New Class of Spatial Statistical Model for Data on Stream Networks: Overview and Applications

Jay Ver Hoef
Erin Peterson
Dan Isaak
Spatial Statistical Models for Stream Networks

Spatial statistical models that use flow and stream distance

Jay M. Ver Hoef · Erin Peterson · David Theobald

Geostatistical modelling on stream networks: developing valid covariance matrices based on hydrologic distance and stream flow

Erin E. Peterson, David M. Theobald and Jay M. Ver Hoef


Authors:
David M. Theobald
John B. Norman
E. Peterson
S. Ferraz
A. Wade
M.R. Sherburne

Spatial modelling and prediction on river networks: up model, down model or hybrid?

Vincent Garreta1, Pascal Monestiez2 and Jay M. Ver Hoef3

1CEREGE, UMR 6635, CNRS, Université Aix-Marseille, Européole de l’Arbois, 13545 Aix-en-Provence, France
2INRA, Unité de Biostatistique et Processus spatiaux, Domaine St Paul, Site Agroparc, 84914 Avignon Cedex 9, France
3NOAA National Marine Mammal Lab, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115, USA
Examples of Autocorrelated Data on Stream Networks

Wenger et al. 2011
10,000 fish sample locations

Garreta et al. 2009
187 nitrate sample locations

Isaak et al. 2010
780 temperature locations

Observations are not independent
Spatial Statistical Models are Dot Connectors

Valid interpolation on networks & aggregation of datasets

Advantages:
- flexible & valid autocovariance structures that accommodate network topology
- weighting by stream size
- improved predictive ability & parameter estimates relative to non-spatial models

Peterson et al. 2006; Ver Hoef et al. 2006; Ver Hoef and Peterson 2010
Spatial Statistical Network Models Work the Way that Streams Do

Gradual trends within networks...

...but also changes at tributary confluences

...& are significantly better mousetraps
Stream Models are Generalizable

Statistical stream models

Response Metrics
- Gaussian
- Poisson
- Binomial

Genetic Attributes

Distribution & abundance

Water Quality Parameters

Stream Temperature
Stream Temperature Database
14 year period (1993 – 2006)
780 observations
518 unique locations

Watershed Characteristics
Elevation range 900 – 3300 m
Fish bearing streams ~2,500 km
Watershed area = 6,900 km²

An Example in the Boise River Basin
Boise River Temperature Model

Non-spatial Stream Temp =
- 0.0064*Elevation (m)
+ 0.0104*Radiation
+ 0.39*AirTemp (°C)
- 0.17*Flow (m³/s)

Parameter estimates are different because of autocorrelation in database

Spatial Stream Temp =
- 0.0045*Elevation (m)
+ 0.0085*Radiation
+ 0.48*AirTemp (°C)
- 0.11*Flow (m³/s)

Isaak et al. 2010. Ecol. Apps. 20:1350-1371
A Kriged River Temperature Map

2006 Mean Summer Temperatures
Spatial estimates are less biased & more precise
Spatial Variation in Prediction Precision
Models Describe Autocorrelation Distances

Distance between samples (km)

Summer Stream Temperature

Inverse Similarity

Redundant information

Efficient Monitoring Designs

Too many...

Too few...

Just right
Autocorrelation Type & Distance Depends on Network Topology

Are sites flow connected or unconnected?
Autocovariance Structures for Stream Models

Sample size & computational requirements

Minimum sample size \( \sim n \geq 50 / 100 \)
- more parameters with autocovariance
- spatial clustering needed

Maximum sample size \( \sim n < 10,000 \)
- inversion of \( n \times n \) matrix
a BIG DATA challenge
>45,000,000 hourly records
>15,000 unique stream sites

>60 agencies
$10,000,000
Regional Temperature Model

Accurate temperature models

Cross-jurisdictional “maps” of stream temperatures

Consistent planning across 350,000 stream kilometers

NorWeST

Stream Temp

$y = 0.68x + 3.82$

$y = 0.55x + 7.79$

$y = 0.86x + 2.43$

$y = 0.93x + 0.830$

Predicted (°C)

Observed (°C)

Replace with validation data from 2007
Example: SpoKoot River Basins

- 5,482 August means
- 2,185 stream sites
- 19 climate summers

Temperature sites:
- Flathead R.
- Bitteroot R.
- Kootenai R.

55,000 stream km
Example: SpoKoot River Basins

Data extracted from NorWeST

- 5,482 August means
- 2,185 stream sites
- 19 climate summers

55,000 stream km

Temperature site

SpoKoot River Basins
SpoKoot River Temp Model

\[ n = 5,482 \]

Covariate Predictors

1. Elevation (m)
2. Canopy (%)
3. Stream slope (%)
4. Ave Precipitation (mm)
5. Latitude (km)
6. Lakes upstream (%)
7. Baseflow Index
8. Watershed size (km\(^2\))
9. Discharge (m\(^3\)/s)
   - **USGS gage data**
10. Air Temperature (°C)
   - **RegCM3 NCEP reanalysis**
   - **Hostetler et al. 2011**

\[ r^2 = 0.90; \ RMSE = 0.97°C \]
Kriged Prediction Map of Climate Scenario
1993-2011 mean August temperatures

20x larger than original Boise model

1 kilometer resolution
55,000 stream kilometers
Better Understanding & Prediction for Streams

New relationships described

Old relationships tested

Refined

Rejected

Predictor

Response
Key References – Theory


Key References - Applications


SSN (Spatial Stream Networks) R Package on CRAN

SSN: Spatial Modeling on Stream Networks

Geostatistical modeling for data on stream networks, including models based on in-stream models, including covariates, can be fit with ML or REML. Mapping and other graphical functions.

Version: 1.1
Depends: R (≥ 2.10), methods, maptools, RSQLite, igraph (≥ 0.6), MASS, sp, BH
LinkingTo: BH
Published: 2013-04-06
Author: Jay Ver Hoef and Erin Peterson
Maintainer: Jay Ver Hoef <ver.hoef at SpatialStreamNetworks.com>
License: GPL-2
NeedsCompilation: yes
Citation: SSN citation info
CRAN checks: SSN results

Downloads:
Package source: SSN_1.1.tar.gz
MacOS X binary: SSN_1.1.tgz
Windows binary: SSN_1.1.zip
Reference manual: SSN.pdf
Vignettes: Model Introduction and SSN User Manual
GIS Information and STARS User manual
Old sources: SSN archive

Related packages
SSN/STARS Website

FreeWare Tools, Example Datasets, & Applications

Google “SSN/STARS”

Analytical Stream Ecosystem is Growing

Modelling dendritic ecological networks in space: an integrated network perspective

A Moving Average Approach for Spatial Statistical Models of Stream Networks

STARS: An ArcGIS toolset used to calculate the spatial data needed to fit spatial statistical models to stream network data
Digital Preprocessed Stream Networks also Available Through SSN/STARS

GIS infrastructure

~350,000 Stream kilometers

Spatial models

NorWeST

Stream Temp

Just need spatial stream datasets
User Community is Growing
>7,000 Visits to SSN/STARS Website in last 9 months

VHP models
User Community is Growing

>7,000 Visits to SSN/STARS Website in last 9 months

VHP models

Global
The End