SPMODEL: SPATIAL MODELING IN **R**

Michael Dumelle *
United States
Environmental Protection Agency
200 SW 35th St, Corvallis, OR, 97333
Dumelle.Michael@epa.gov

Matt Higham
Department of Math, Computer Science, and Statistics
St. Lawrence University
23 Romoda Drive, Canton, NY, 13617
mhigham@stlawu.edu

Jay M. Ver Hoef
National Oceanic and Atmospheric Administration
Alaska Fisheries Science Center
Marine Mammal Laboratory, Seattle, WA, 98115
jay.verhoef@noaa.gov

April 23, 2022

Abstract

Enter the text of your abstract here.

Keywords Spatial covariance · Linear Model · Autoregressive model

1 Introduction

Here we describe the general role of spatial modeling, discuss existing software, and argue spmodel is a valuable contribution.

2 Background and Usage

2.1 Spatial Linear Models

```
R> spmod <- splm(y ~ x, exdata, "exponential", xcoord, ycoord)
R> summary(spmod)


Call:
splm(formula = y ~ x, data = exdata, spcov_type = "exponential",
    xcoord = xcoord, ycoord = ycoord)

Residuals:
     Min      1Q   Median      3Q      Max
-2.17030 -0.46466  0.07753  0.56000  2.41177

Coefficients (fixed):
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.39465    0.38302  -1.030    0.303
x           -0.07311    0.07374  -0.991    0.322
```

*Corresponding Author

```
Pseudo R-squared: 0.009929

Coefficients (spatial covariance):
    de     ie  range
0.6561 0.3874 1.3258

Spatial covariance type: exponential
```
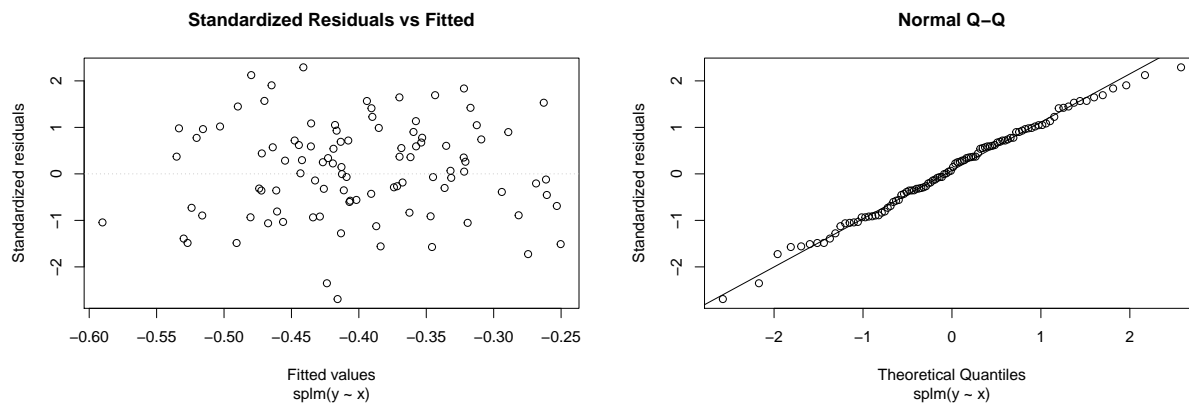
- REML citations (Patterson and Thompson 1971; Harville 1977; Wolfinger, Tobias, and Sall 1994)
- SV-WLS citations (Cressie 1985, 1993)
- SV-CL citations (Curriero and Lele 1999)

```
R> plot(spmod, which = c(1, 2))
```


2.1.1 Anisotropy

```
R> spmod <- splm(y ~ x, exdata, "exponential", xcoord, ycoord, anisotropy = TRUE)
```

2.2 Prediction

```
R> spmod_preds <- predict(spmod, newexdata, interval = "prediction")
```

2.3 Model-fit Statistics

2.4 broom functions

```
R> tidy(spmod)

# A tibble: 2 x 5
  term        estimate std.error statistic p.value
  <chr>          <dbl>     <dbl>     <dbl>   <dbl>
1 (Intercept)  -0.391     0.364     -1.07    0.283
2 x            -0.0713    0.0738    -0.966   0.334

R> glance(spmod)

# A tibble: 1 x 9
      n     p  npar value   AIC  AICc logLik deviance pseudo.r.squared
  <int> <dbl> <dbl> <dbl> <dbl> <dbl>  <dbl>    <dbl>            <dbl>
1   100     2     5  244.  256.  257.  -122.       98          0.00943

R> augment(spmod)
```

```
# A tibble: 100 x 11
        x       y   xcoord  ycoord group subgroup .fitted  .resid    .hat .cooksd .std.resid
    <dbl>   <dbl>    <dbl>   <dbl> <fct> <fct>      <dbl>   <dbl>   <dbl>   <dbl>      <dbl>
 1 -1.12    0.602  -0.783   0.0723 1     1         -0.310  0.913  0.135  0.0642       1.05
 2 -0.339   1.40   -0.847  -0.805  1     2         -0.366  1.77   0.0683 0.0882       1.67
 3  2.67   -0.976  -0.0836  0.382  1     1         -0.581 -0.395  0.115  0.0500      -0.991
 4  0.842  -0.981  -0.313   0.0923 1     2         -0.451 -0.531  0.0150 0.00740     -1.00
 5 -0.283  -0.724   0.521   0.477  1     1         -0.370 -0.353  0.0483 0.00183     -0.282
 6  0.438  -0.182   1.45   -0.703  1     2         -0.422  0.239  0.0644 0.00245      0.285
 7  0.650   0.109   0.433  -1.18   1     1         -0.437  0.546  0.0287 0.00158      0.337
 8 -0.315  -0.342  -0.0726 -2.59   1     2         -0.368  0.0262 0.0641 0.00238     -0.282
 9  0.289  -2.59    1.52   -1.11   1     1         -0.411 -2.17   0.0217 0.0771      -2.70
10  0.315   0.383   1.21    1.05   1     2         -0.413  0.796  0.0452 0.0236       1.05
# ... with 90 more rows

R> augment(spmod, newdata = newexdata)

# A tibble: 10 x 6
       x xcoord ycoord group subgroup .fitted
   <dbl>  <dbl>  <dbl> <fct> <fct>      <dbl>
 1  0.701 -0.857 -1.03  4     2          0.646
 2 -1.45  -0.396  1.92  1     2         -0.954
 3 -1.29  -2.05  -0.238 3     1         -0.233
 4 -0.361  0.350 -0.243 2     2         -0.322
 5 -0.802 -1.76  -1.98  1     2         -0.442
 6  1.45  -1.04  -2.19  2     1         -0.504
 7  0.218 -0.125  2.22  2     2         -0.972
 8  0.480 -1.08   0.426 1     1         -0.0898
 9  0.908  0.378 -2.29  4     1         -0.592
10 -0.883  1.57  -0.601 2     1         -0.173
```

2.5 Spatial Autoregressive Models

```
R> spmod <- spautor(y ~ x, exdata_poly, "car")
```

2.6 Neighborhood Indexing for Big Data

```
R> spmod <- splm(y ~ x, exdata, "exponential", xcoord, ycoord, local = TRUE)
```

2.6.1 The local list

```
R> spmod <- splm(y ~ x, exdata, "exponential", xcoord, ycoord, local = list(parallel = TRUE))
```

2.7 Random Effects

```
R> spmod <- splm(y ~ x, exdata, "exponential", xcoord, ycoord, random = ~ group)
```

2.8 Partition Factors

```
R> spmod <- splm(y ~ x, exdata, "exponential", xcoord, ycoord, partition_factor = ~ group)
```

2.9 Initial Values and Known Values

```
R> spcov_params_init <- spcov_initial("exponential", de = 1, ie = 0, known = "ie")
R> spmod <- splm(
+    y ~ x,
+    exdata,
+    spcov_initial = spcov_params_init,
+    xcoord = xcoord,
```

```
+    ycoord = ycoord
+ )
```

2.10 Simulating Gaussian Random Variables

```
R> spcov_params_val <- spcov_params("exponential", de = 1, ie = 1, range = 2)
R> var <- sprnorm(spcov_params_val, data = exdata, xcoord = xcoord, ycoord = ycoord)
```

3 Discussion

Data and Code Availability

Acknowledgements

References

Cressie, Noel. 1985. "Fitting Variogram Models by Weighted Least Squares." *Journal of the International Association for Mathematical Geology* 17 (5): 563–86.

———. 1993. *Statistics for Spatial Data.* John Wiley & Sons.

Curriero, Frank C, and Subhash Lele. 1999. "A Composite Likelihood Approach to Semivariogram Estimation." *Journal of Agricultural, Biological, and Environmental Statistics*, 9–28.

Harville, David A. 1977. "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems." *Journal of the American Statistical Association* 72 (358): 320–38.

Patterson, H Desmond, and Robin Thompson. 1971. "Recovery of Inter-Block Information When Block Sizes Are Unequal." *Biometrika* 58 (3): 545–54.

Wolfinger, Russ, Randy Tobias, and John Sall. 1994. "Computing Gaussian Likelihoods and Their Derivatives for General Linear Mixed Models." *SIAM Journal on Scientific Computing* 15 (6): 1294–1310.