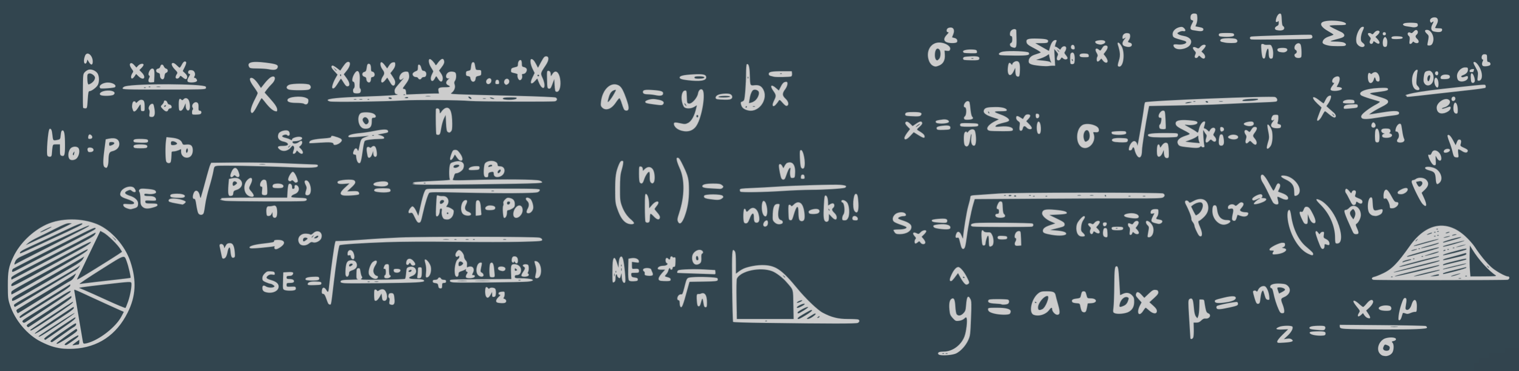
Variogrammes

Ça se discute

A. Boyé



Spatial (or temporal) correlation may be described by [1]

- **Correlograms**

A correlogram is a graph in which *spatial correlation* values are plotted, on the ordinate, against *distance classes* among sites on the abscissa.

- **Variograms** (semi-variograms)

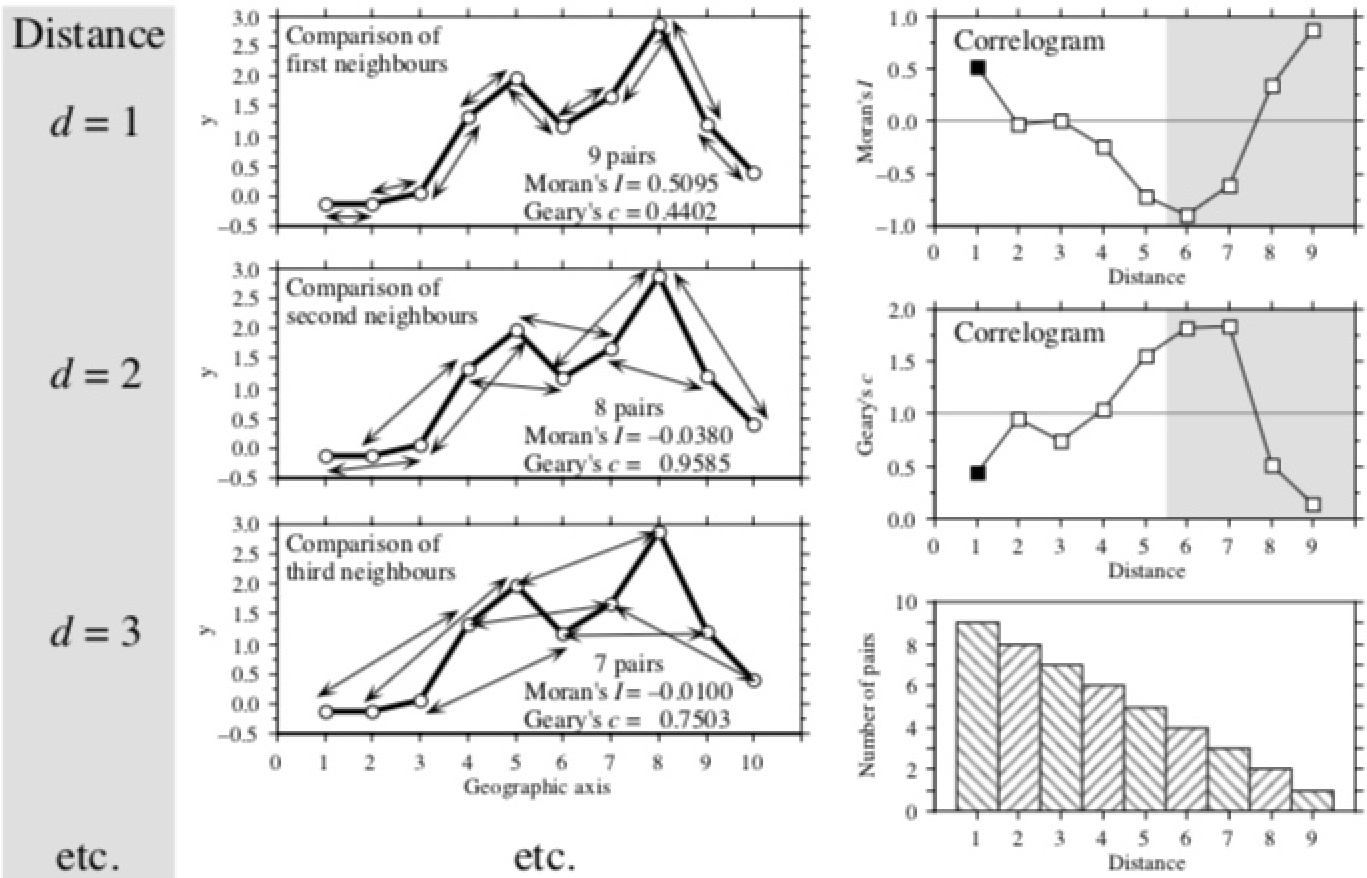
A variogram is a graph in which *semi-variance* is plotted, on the ordinate, against *distance classes* among sites on the abscissa

- **Periodograms** (structure assumed to consist of *a combination of sine waves*)

Fit sines and cosines of various periods, one period at a time, and determine the proportion of the series' variance (R^2) explained by each period. The abscissa is either a period or its inverse, a frequency; the ordinate is the proportion of variance explained.

[1] Legendre, P. and Legendre, L. (2012). *Numerical Ecology, 3rd English edition.* Elsevier.

Principle



Distance
d = 1
d = 2
d = 3
etc.
Comparison of first neighbours
9 pairs
Moran's I = 0.5095
Geary's c = 0.4402
Comparison of second neighbours
8 pairs
Moran's I = –0.0380
Geary's c = 0.9585
Comparison of third neighbours
7 pairs
Moran's I = –0.0100
Geary's c = 0.7503
Geographic axis
etc.
Correlogram
Moran's I
Distance
Correlogram
Geary's c
Distance
Number of pairs
Distance

Correlograms

$$\text{Moran's } I: \quad I(d) = \frac{\frac{1}{W}\sum_{h=1}^{n}\sum_{i=1}^{n} w_{hi}(y_h-\bar{y})(y_i-\bar{y})}{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2} \quad \text{for } h \neq i \qquad (13.1)$$

$$\text{Geary's } c: \quad c(d) = \frac{\frac{1}{2W}\sum_{h=1}^{n}\sum_{i=1}^{n} w_{hi}(y_h-y_i)^2}{\frac{1}{(n-1)}\sum_{i=1}^{n}(y_i-\bar{y})^2} \quad \text{for } h \neq i \qquad (13.2)$$

- **Moran's I**

Moran's I related to the *Pearson correlation* coefficient. *Usually* takes values between −1 and +1

Numerator : *covariance*;

Denominator : maximum-likelihood estimator of the *variance*

- **Geary's c**

Geary's c coefficient is a *distance-type* function : varies from 0 to some unspecified value larger than 1.

Numerator : sums the squared differences

Varies as the reverse of a Moran's I correlogram

Correlograms

$$\text{Moran's } I: \qquad I(d) = \frac{\frac{1}{W}\sum_{h=1}^{n}\sum_{i=1}^{n} w_{hi}(y_h - \bar{y})(y_i - \bar{y})}{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad \text{for } h \neq i \qquad \textbf{(13.1)}$$

$$\text{Geary's } c: \qquad c(d) = \frac{\frac{1}{2W}\sum_{h=1}^{n}\sum_{i=1}^{n} w_{hi}(y_h - y_i)^2}{\frac{1}{(n-1)}\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad \text{for } h \neq i \qquad \textbf{(13.2)}$$

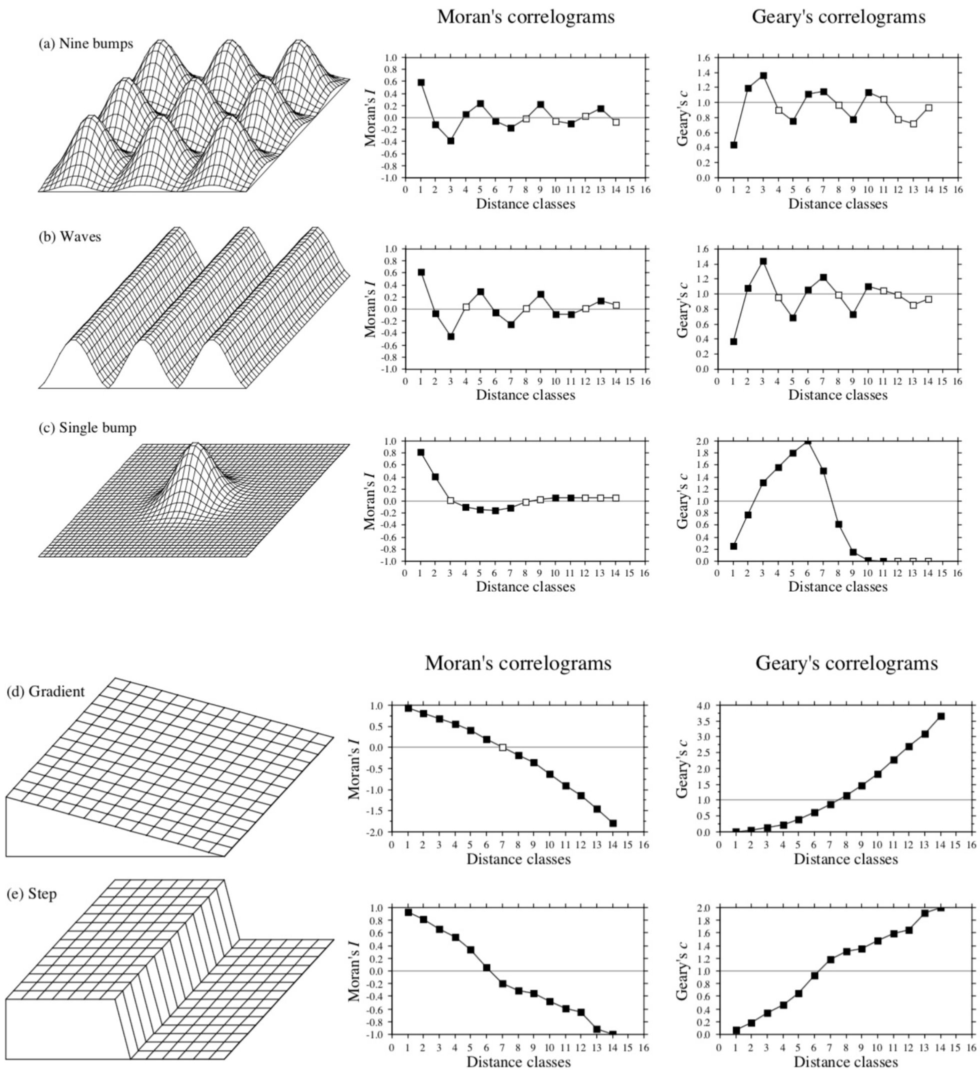
Positive spatial correlation : positive values of I.

Negative correlation : negative values.

$\Rightarrow$ **Strong spatial correlation produces high values of I and low values of c**

Sensitive to extreme values and asymmetry in distribution : normalize data

Absence of auto-correlation : Moran’s I = 0 & Geary’s c = 1



Correlograms

Moran's I: $$I(d) = \frac{\frac{1}{W}\sum_{h=1}^{n}\sum_{i=1}^{n} w_{hi}(y_h - \bar{y})(y_i - \bar{y})}{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad \text{for } h \neq i \qquad \textbf{(13.1)}$$

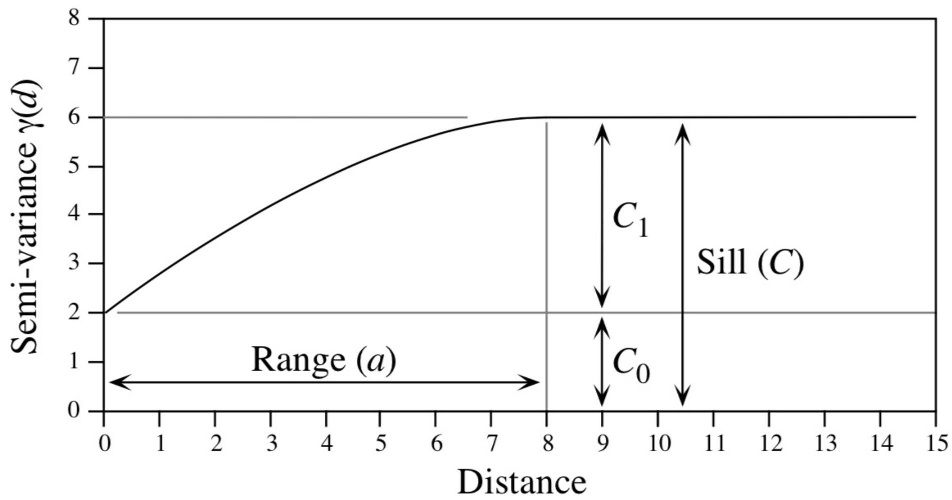
Geary's c: $$c(d) = \frac{\frac{1}{2W}\sum_{h=1}^{n}\sum_{i=1}^{n} w_{hi}(y_h - y_i)^2}{\frac{1}{(n-1)}\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad \text{for } h \neq i \qquad \textbf{(13.2)}$$

Semi-variograms

$$\gamma(d) = \frac{1}{2W(d)}\sum_{h=1}^{n}\sum_{i=1}^{n} w_{hi}(y_h - y_i)^2 \quad \text{for } h \neq i \qquad \textbf{(13.9)}$$

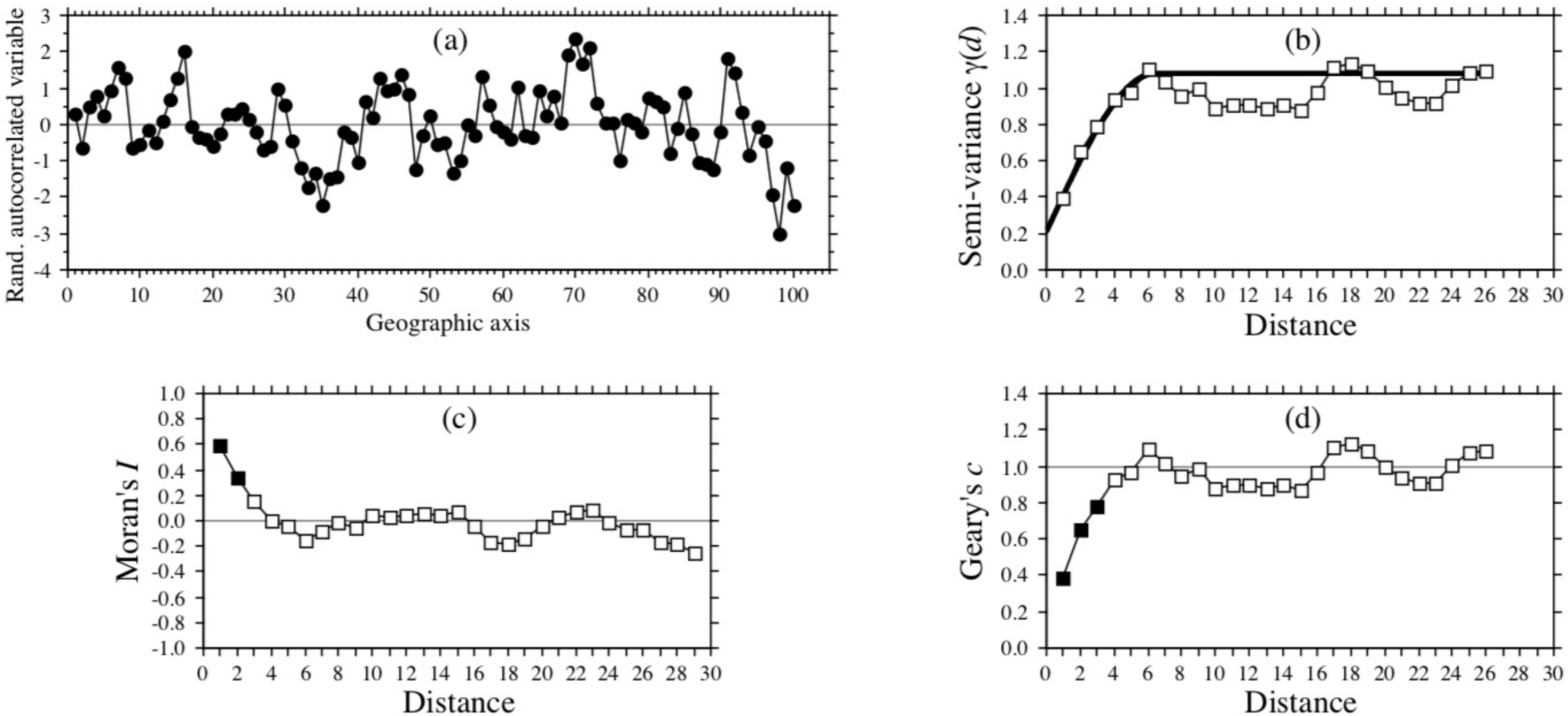
$\Rightarrow$ **Numerator of Geary's c**

Semi-variograms



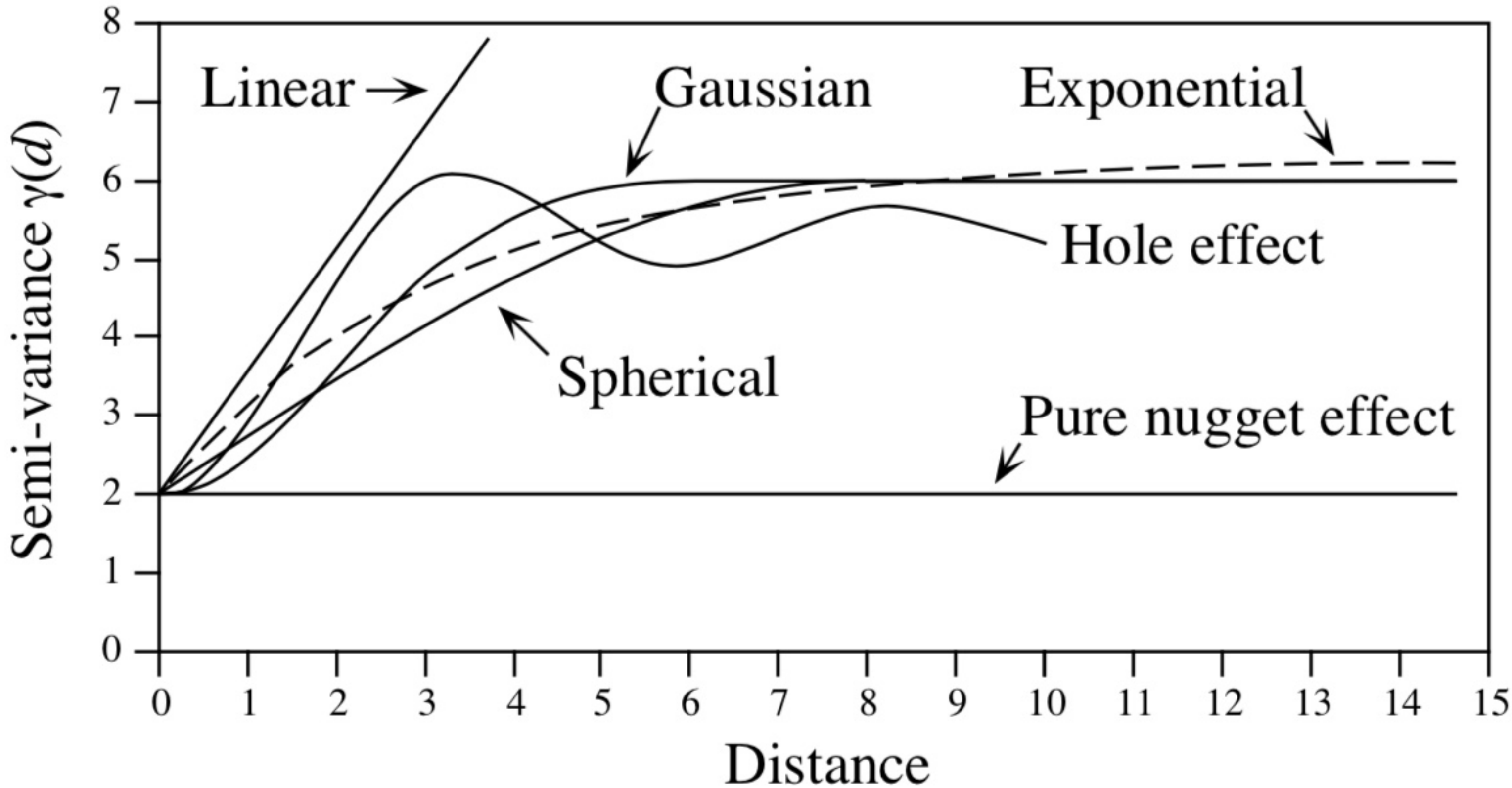
- *sill* : variance of the variable
- *range* : beyond that distance, the sampling units are not spatially correlated
- C_0 (*nugget effect*) : local variation occurring at scales finer than the sampling interval
    - ex : sampling error, fine-scale spatial variability, and measurement error
- C_1 : spatially structured component

Example



(a)
Rand. autocorrelated variable
Geographic axis
(b)
Semi-variance γ(d)
Distance
(c)
Moran's I
Distance
(d)
Geary's c
Distance

Fit models to the sample variogram (kriging)



Computation in R

Autocorrelation

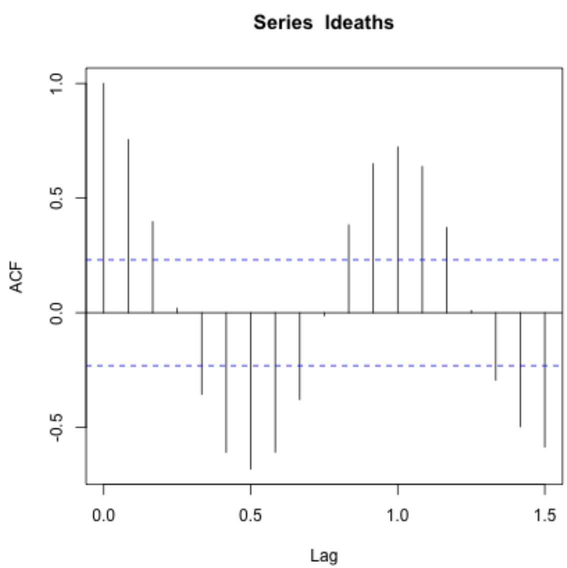
```
acf(x, lag.max = NULL, type = c("correlation", "covariance",
    "partial"), plot = TRUE, na.action = na.fail...)
```

Cross-correlation/cross-variogram

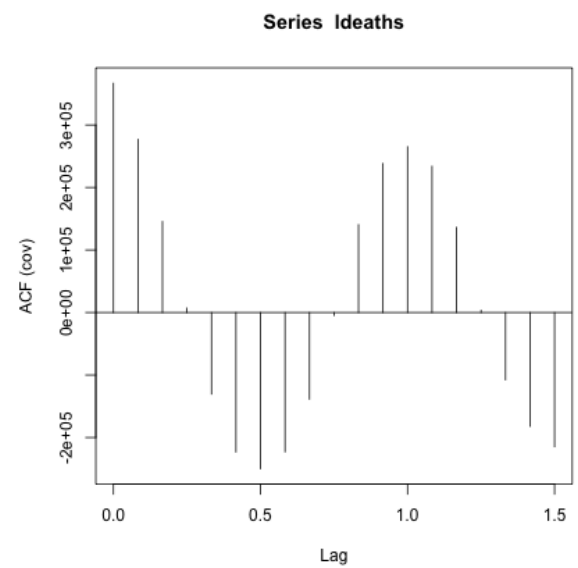
```
ccf(x, y, lag.max = NULL, type = c("correlation", "covariance"),
    plot = TRUE, na.action = na.fail, ...)
```

Monthly Deaths from Lung Diseases in the UK, 1974–1979

```r
acf(ldeaths,
 type = "correlation")
```



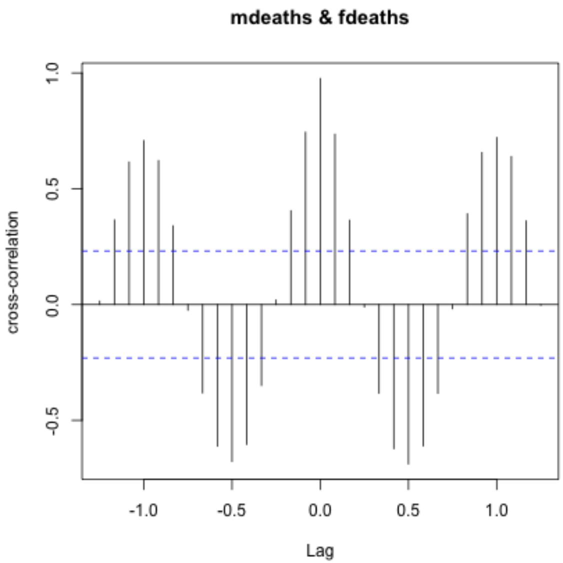

```r
acf(ldeaths,
 type = "covariance")
```

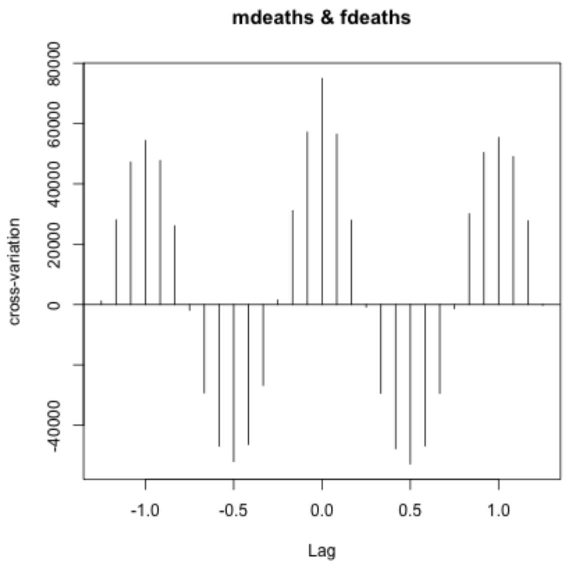


Covariation male/female

```
ccf(mdeaths, fdeaths,
  type = "correlation",
 ylab="cross-correlation")
```

```
ccf(mdeaths, fdeaths,
  type = "covariance",
 ylab="cross-variation")
```





Another example from marine systems [2]

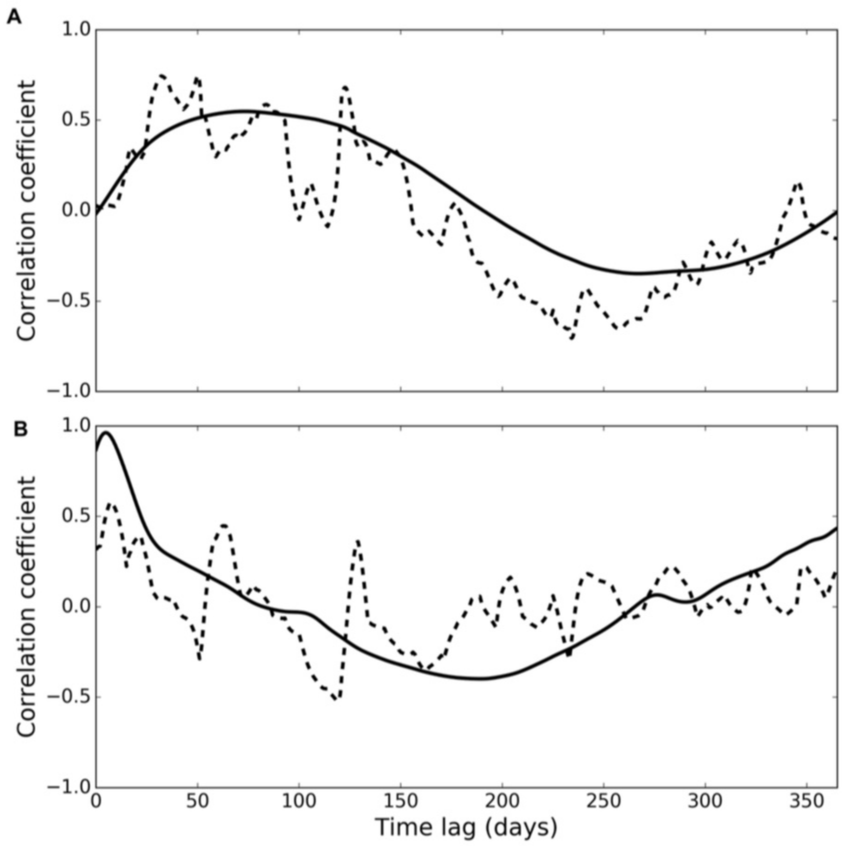


FIGURE 5 | Cross-correlation of phytoplankton with observed (dashed line) and modelled (solid line) deposit feeders **(A)** and suspension feeders **(B)**.

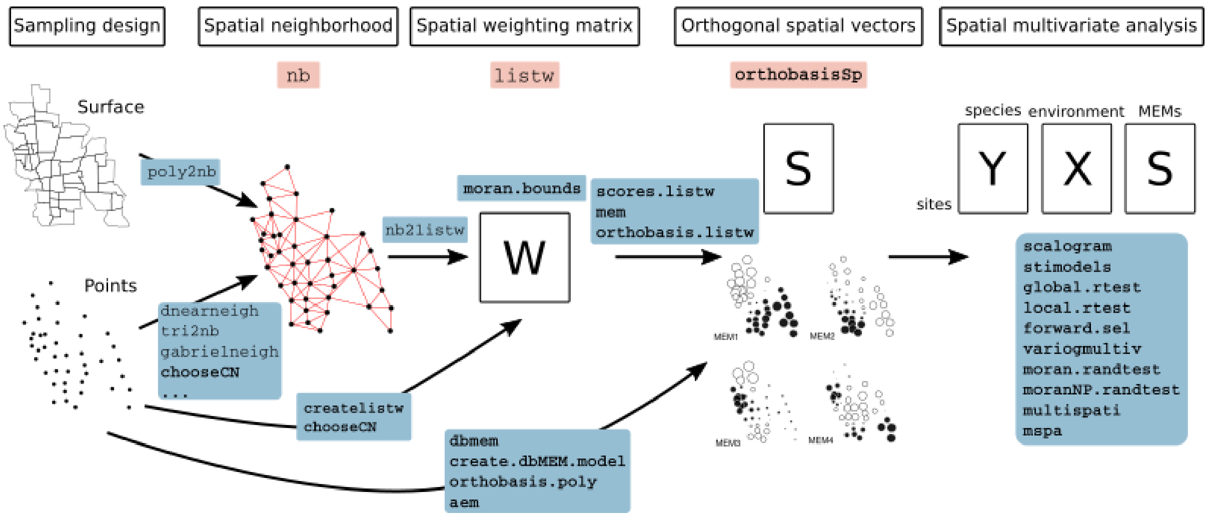
[2] Lessin, G., Bruggeman, J., McNeill, C. L., and Widdicombe, S. (2019). Time scales of benthic macrofaunal response to pelagic production differ between major feeding groups. *Frontiers in Marine Science*, 6:15.

Other approaches (MEM)

Principle

MEM[3] (Moran's eigenvector maps): spatial variables representing structures of all relevent scales

Geographic distances among sites or *spatial weighting matrix* $\Rightarrow$ **spatial explanatory variables**



[3] Dray, S., Legendre, P., and Peres-Neto, P. R. (2006). Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecological Modelling, 196(3):483 – 493.

Examples [4]

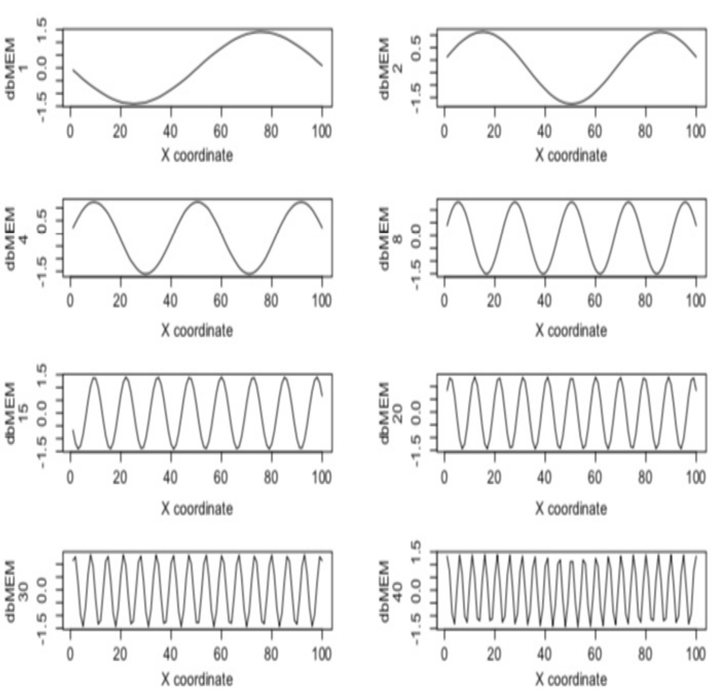


Fig. 7.3 Some of the 49 dbMEM variables with positive eigenvalues built from a transect with 100 equispaced points

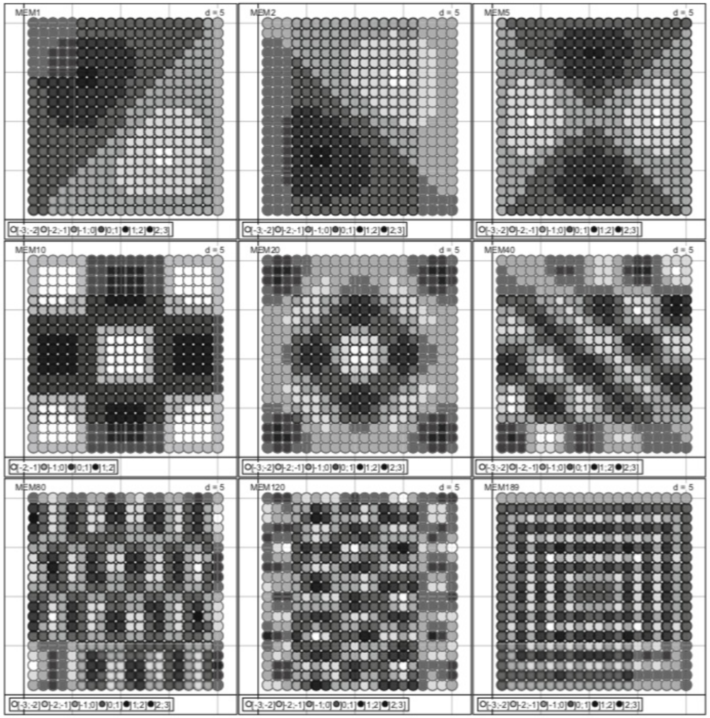


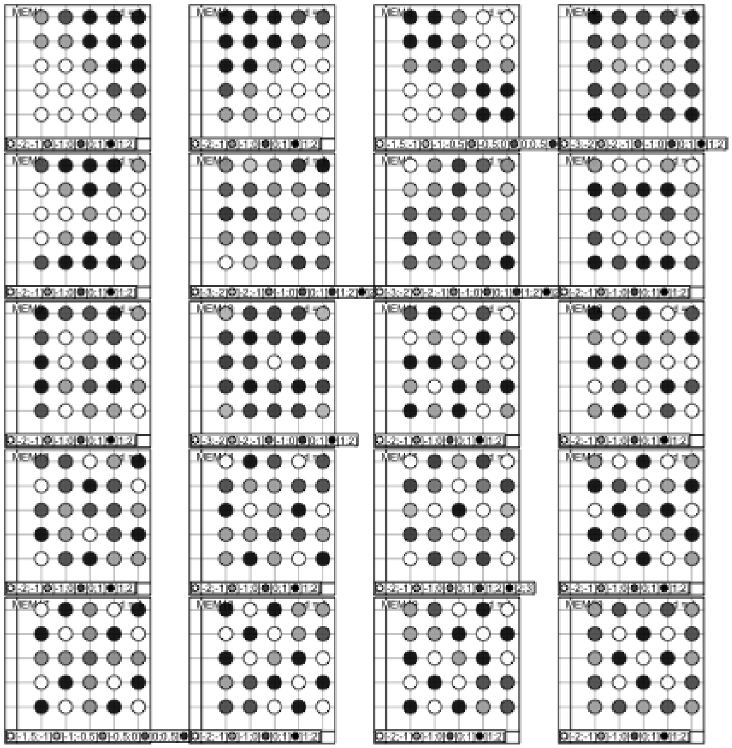
Fig. 7.4 Some of the 189 dbMEM variables with positive eigenvalues built from a grid comprising 20 by 20 equispaced points

[4] Borcard, D., Gillet, F., & Legendre, P. (2018). Numerical ecology with R. 2nd edition. Springer.

MEM and dbMEM can be computed using adespatial

```r
library(adespatial)
library(adegraphics)

# Generate grid point coordinates
xygrid2 <- expand.grid(1:5, 1:5)

# Creation of the dbMEM eigenfunctions w
xygrid2.dbmem.tmp <-
  dbmem(xygrid2,MEM.autocor="non-null")

xygrid2.dbmem <-
  as.data.frame(xygrid2.dbmem.tmp)

# Plot some dbMEM variables using s.valu
somedbmem2 <- c(1:20)
s.value(xygrid2,
        xygrid2.dbmem[,somedbmem2],
  method = "color",
  symbol = "circle",
  ppoints.cex = 0.5
)
```

```
# Information: Square regular grid; multi
```



Analyse de codépendance

Ecology, 91(10), 2010, pp. 2952–2964

Multiscale codependence analysis: an integrated approach to analyze relationships across scales

Guillaume Guénard,[1,3] Pierre Legendre,[1] Daniel Boisclair,[1] and Martin Bilodeau[2]

[1]*Département de Sciences Biologiques, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, Québec H3C 3J7 Canada*
[2]*Département de Mathématiques et de Statistique, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, Québec H3C 3J7 Canada*

Analyse de codépendance

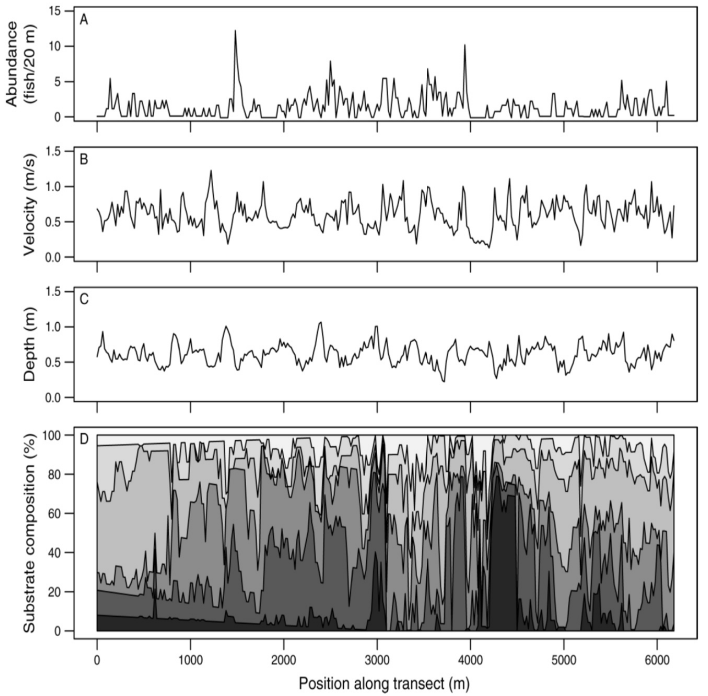


FIG. 4. (A) Abundance of Atlantic salmon parr after the effect of among-day temperature differences and linear trend were removed by regression. Also shown are (B) flow velocity, (C) channel depth, and (D) substrate composition after removal of linear trends. The six levels of shading represent the contributions of the six grain size classes of the river bed, from fine substrate (light) to metric boulder (dark).

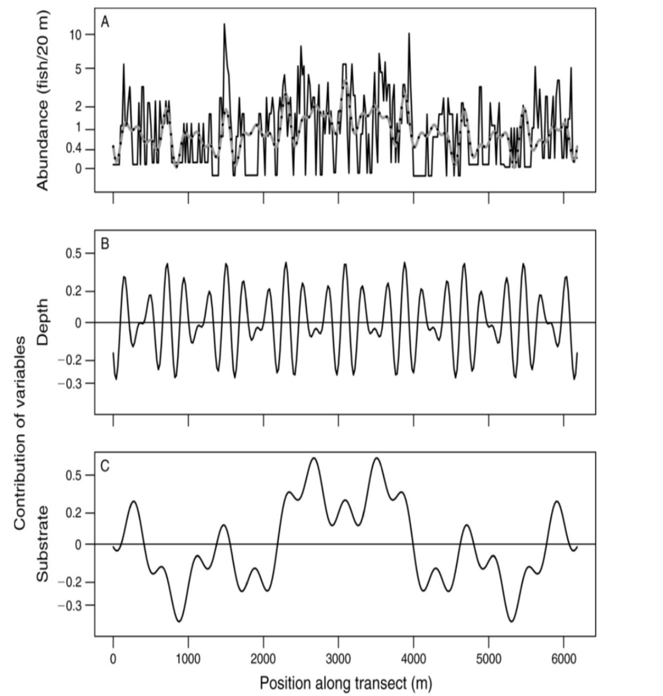


FIG. 5. (A) Observed (solid) and fitted (dotted) parr abundances, on a $\log_e(y+1)$ scale, along the river transect; (B) the fraction of parr abundance explained by channel depth; and (C) substrate composition obtained from Eq. 9. The model explains 18.9% of the $\log(x+1)$-transformed variation in parr abundance.